THE UNIVERSITY OF CHICAGO

University of Chicago
Chicago

PhD thesis

A Mathematical Theory of Central Nervous Activity

Author:
Jack D. Cowan

First Examiner:
Prof. Dr. Raymond Beurle

Advisor:
Prof. Dr. Dennis Gábor

Department:
Electrical Engineering

Submitted: July 15th, 1967

Declaration of Authorship

I declare that the work presented here is original and the result of my own investigations. Formulations and ideas taken from other sources are cited as such. It has not been submitted, either in part or whole, for a degree at this or any other university.

Location, Date

Signature

Abstract

A mathematical model is developed of the activity of networks of model nerve-cells or neurons: A nonlinear delay operator is introduced to represent the transfer characteristics of neurons. This operator is a continuous function that represents the mean neuronal response to stimulating currents. The mathematics used is not the Boolean algebra of switching circuits but Differential equations. The dynamics is examined of certain model brain circuits, some of which are shown to exhibit undamped responses to stimuli, but only for a few levels of activity. The techniques of Hamiltonian mechanics and of Gibbs'ian statistical mechanics are used to connect these models with experimental data. A preliminary explanation is given for the existence of preferred states found in the firing patterns of neurons in animal nervous systems.

Contents

1. Introduction **10**

2. Formulation of the Neuronal Equation **19**

3. Neuronal Oscillations **25**

4. Corticothalamic Networks **31**

5. Statistical Mechanics of Systems of Neuronal Oscillators **35**
5.1. Calculation of "canonical" averages 37

6. Some Applications to Experimental Data **55**

7. Summary. Relation to other Work. Concluding Remarks **61**

8. Acknowledgements **64**

A. Redundant Automata as Models of Neuron Assemblies **65**
A.1. Description of Problem 65
A.2. Acknowledgement 70

B. Neuronal Nets **71**
B.1. Introduction 71
B.2. The Neuron 73
B.3. Neuronal Nets 78
B.3.1. The Electroencephalogram 78
B.3.2. Electrocorticograms and Evoked Responses 80
B.3.3. Microelectrode Recordings and Slow Potentials 84
B.3.4. Functional Aspects of Nets 85
B.4. Discussion 90
B.5. Acknowledgement 92

C. Figures **93**

References **148**

List of Figures

C.1. The model neuron 94
C.2. Two current impulses – sub-synaptic region and "Smeared" 95
C.3. The logistic function 95
C.4. $\ln \frac{X}{1-X}$ vs. X 96
C.5. Postulated Corticothalamic circuits 97
C.6. Interneuron circuits 98
C.7. Cortical anatomy 99
C.8. Recurrent collateral and afferent excitation and inhibition 100
C.9. Interactions of cortical and thalamic neurons 101
C.10. The Beta density function 102
C.11. Sensitivity oscillations 103
C.12. Frequency, discharge and intracellular current of a cell 104
C.13. $\ln \frac{X}{1-X}$ vs. I estimated from extracellular experiments 104
C.14. Expectation Densities for a thalamic-somatic sensory neuron 105
C.15. Frequency of oscillations in firing rate of neurons 106
C.16. Temperature Histograms 107
C.17. Cumulative Interval Distributions for two cortical cells 108
C.18. A communication system 108
C.19. A computation system 108
C.20. Error probability vs. Rate of transfer of information 109
C.21. Uncoded automaton 110
C.22. Coded automaton 111
C.23. Cerebral Hemispheres 112
C.24. Axonal Bundles 114
C.25. A Neuron 115
C.26. Neuronal membrane – an equivalent circuit 116
C.27. Some typical responses of the equivalent circuit 117
C.28. The effects of distributed parameters 118
C.29. IS-membrane – an equivalent circuit 118
C.30. Analogue computer solutions of the Hodgkin–Huxley equations . . . 119
C.31. Rate-intensity curves of neuronal spike emission 120
C.32. Resting EEG of a normal adult 122
C.33. Autocorrelogram of resting EEG of normal adult 122
C.34. Probability density of amplitudes of EEG 124
C.35. Changes in the normal EEG during alertness, drowsiness and sleep 125
C.36. Abnormal EEGs 126

C.37.The cortical frequency spectrum 126
C.38.Evoked potential in the visual cortex, cat 127
C.39.Response to flash in man 127
C.40.Cortico-subcortical circuits 128
C.41.Spike patterns and EEG 128
C.42.Correlation between PSTHs and mean slow potential 129
C.43."ON", "OFF", and "ON OFF" units 130
C.44.Discharge produced by central spot, inhibited 131
C.45.Mutual inhibition of two receptor units 132
C.46.Retinal anatomy 133
C.47.A simple and a complex cell 134
C.48.A hypercomplex cell 135

List of Tables

A.1. Possible patterns of zero-one triplets 67

1. Introduction

The last quarter of a century has seen the emergence and development of cybernetics, the science of information and control. As a result a new language has evolved out of the continued efforts of scientists and engineers to make precise such concepts as 'communication channel', 'automaton', and 'control system'. This language has been widely used in those sciences which have to deal with the interaction of animals and machines with their environments. One science that has drawn heavily on this language is neurology. Cybernetic concepts have been used in many different ways to investigate how the brains of men and animals might work.
Perhaps the earliest example of the use of this language in neurology is to be found in the 1943 publication by W. S. McCulloch and W. H. Pitts of "A logical calculus of the ideas immanent in nervous activity" [104]. In this paper the concept formal neuron' was introduced, essentially an abstraction from then current details of neuronal operation. These formal neurons operated upon and emitted, at specified times, binary "all-or-none" signals. Their junctions, called "synapses", were either excitatory or else totally inhibitory. They functioned by computing the algebraic sum of the values of their inputs, 1 and 0, subsequently emitting an output signal if and only if this sum exceeded a certain specified threshold. Formal neurons could be made to represent the elementary operations of two-valued logic, and a fortiori, formal neuronal networks to represent complicated logical formulas. The main result of the paper (subsequently proved by S. C. Kleene [87]) is that any mode of the behaviour of brains that can be finitely expressed with linguistic precision (corresponding to some well-formed formula of the calculus of first-order predicates) can be realized by a finite network of these formal neurons. To the extent to which the concept 'formal neuronal network' symbolizes the process whereby brains respond to and represent stimuli, the McCulloch-Pitts theory is similar to that of K. J. Craik's regarding the nature and functioning of neuronal networks ([31]). A corollary of the McCulloch-Pitts theorem is also of interest: formal neuronal networks plus potentially infinite auxiliary stores are equivalent to Turing machines. A Turing machine [151] is itself a formalization of, and an abstraction from the processes underlying computing. The equivalence of the two concepts therefore led immediately to the idea that in at least some aspects of signal processing, brains and computers are similar. N. Wiener in his now classic book "Cybernetics" [165] developed this analogy, incorporating chapters on computing machines and the nervous system, on gestalts and universals and on cybernetics and psychopathology.
Any direct analogy between formal neuronal networks and brains requires the

existence in brains of specific circuits, so that there has to be a process of selection of nerve-cells and of the pattern of interconnexion, the anatomy or the "wiring-diagram", and a rejection of all other possibilities. However ablation studies on the brains of many animals [93] indicated that such complex signal processing tasks as visual integration were apparently independent of the specific details of the wiring diagrams of the visual cortex. Even extensive destruction of this tissue did not produce marked disintegration of function, although those activities that persisted were somewhat retrenched. This suggested that perhaps the circuits were redundant throughout areas of nervous tissue and that only gross parameters of the tissue such as the mean number of cells and their interconnexions were reliable measures of the tissues' functioning. Wiener (op.cit.) suggested that nervous tissue might be similar to large telephone exchanges in containing redundant cells and interconnexions, and that messages might be simultaneously transmitted from area to area along many distinct pathways, and might be repeated several times along each pathway, the final decision concerning the validity of signals reaching any area of nervous tissue being determined by some kind of voting procedure. J. von Neumann, [109], in the paper, "Probabilistic logics and the synthesis of reliable organisms from unreliable components" gave the first proof of the existence of designs for the construction of neuronal networks which might survive extensive malfunctions of components, or failures of individual formal neurons and of errors in and damage to their connexions. Von Neumann's designs utilized the replication of many individual circuits, the simultaneous transmission of messages by many circuits, and "majority-voting" circuits to ensure the overall reliability of signal processing and transmission. Von Neumann was not satisfied with the results. He considered his treatment of error to be rather ad hoc, and he suggested that error should be the subject of a thermodynamical theory, as C. E. Shannon had treated the concept of information [139][140]. He was also well aware that formal neuronal networks were digital in their mode of operation, whereas those comprising brains were not, despite their use of pulses in some operating modes (see [135]). S. Winograd and the author extended Shannon's theorem concerning the reliable transmission of messages through a noisy communication channel to include computation, and so provided part of von Neumann's suggested thermodynamical theory of error for formal neuronal networks, [170]. See Appendix A for a survey of some of the details. The resulting design for the construction of formal neuronal networks that function correctly in spite of malfunctions of individual neurons, or of errors in and of damage to wiring diagrams, differs in interesting ways from von Neumann's. The ability of formal neuronal networks to function correctly in spite of vicissitudes depends on their composition by anastomotically redundant circuits. In von Neumann's prospectus this kind of redundancy was sought by incorporating many copies of the one and only one circuit necessary for the representation of the specified functions which are intended to determine the behaviour of the network. In the Winograd-Cowan theory however, the redundancy is obtained not by multiplication of identical circuits, but by replacing the specified

functions by a greater variety of more complicated functions which require circuits containing many more and richer formal neurons and interconnexions than the circuits required for the original specified functions. The key to theory is that the rules governing this replacement constitute error-correcting codes [114]. The network that results from such an encoding is immune to many kinds of error, i.e. it is largely "error-insensitive". The degree of error-insensitivity obtained depends upon the complexity of the requisite behaviour, on the frequency of errors, and on the level of redundancy introduced by whatever code is used. Because the code operates on functions and not on circuits, the redundancy obtained is functional i.e., more functions are represented in the encoded network than in its precursor. Because of the nature of error-correcting codes, any one specified function appears in many of these encoded functions, and any one encoded function is essentially a different mixture of many of the specified functions. It is the multiple representation of a multiplicity of specified functions which leads to error-insensitive operation, rather than the mere replication of circuits. In short it is the diversity of the encoded function that is computed by each individual formal neuron comprising the encoded network which leads to the efficiency of design. So the number of formal neurons required to realize any given mode of behaviour at some level of error-insensitivity, despite probability of malfunction of individual formal neurons and of error in the wiring, is ultimately as small as Shannon's theorem indicates is possible. Naturally the formal neurons required to represent encoded functions are much more complicated than those which would be required to represent only the original specified functions. The extension of Shannon's theorem thus requires that the more complicated formal neurons be no less reliable than the simpler ones, a requirement which is equivalent to assuming that the extra "hardware" required for coding is completely error insensitive, as in Shannon's theorem. The status of "computation capacity", the analogue of Shannon's notion of "channel capacity", depends upon the validity of this assumption. The Winograd-Cowan theorem is in fact like von Neumman's theorem, an existence theorem; but it differs from it in requiring a minimal number of very complicated circuits and formal neurons rather than a large number of simple circuits of threshold formal neurons. Our theory is the other extreme to von Neumann's theory wherein complication is minimised at the cost of increased replication, in that redundancy is minimised at the cost of increased complication, to attain some requisite level of error-insensitivity. Suitable combinations of the two techniques, functional coding followed by the replication of the resulting circuits, plus a randomization of the interconnexions between them, lead to efficient and practical error-insensitive networks.
However these networks can persist only for lifetimes that are limited by the error and failure rates of their components. There is an ageing effect, so that the reliability of function of these networks degrades under the cumulating effects of uncorrected failures. W. H. Pierce [116] has shown how "adaptive" networks may be designed whose lifetimes are substantially longer than those of the above networks. A basic defect of the techniques outlined, in which majority voting effectively

controls errors, is that, since those inputs to a given formal neuron which issue from a failed one will be permanently in error a consistently reliable minority may be outvoted by a consistently unreliable majority. Such a limitation may be overcome by using more complicated formal neurons whose inputs are weighted according to their reliabilities. This requires computation by the recipient formal neuron of these reliabilities. Pierce has proved that such computations can be carried out. If input errors are statistically independent, the input weights i be selected so that the output of the formal neuron is the digit most likely to be correct, on the assumption that what is required of the formal neuron is the representation of the majority function, [105]. One variant of this is particularly interesting. In this, if p_i is the number of coincidences and q_i the number of disagreements between the i-th input to, and the output from a given formal neuron in a cycle of m operations, then the selections $s_i = \log(p_i/q_i)$ lead to the computation of the most likely estimate of the reliabilities of the inputs for the representation of the majority function. There are evidently good grounds for asserting that reliable networks of competent formal neurons can be constructed, using the techniques of functional coding followed by replication of the resulting circuits, together with feedback controlled selection of vote-weights, which would be efficient and long-lasting compared with their components. However the question remains of how such networks are to be constructed. The existence of specific circuits in a network requires the selection of components and wirings. For these redundant networks the problem is crucial; i.e., the amount of selective information required to specify them is very large. Not only do the wirings and thresholds have to be specified but also particular patterns of synaptic interactions. Vote weights do not have to be selected because the process is automatic, but the necessary wirings of formal neurons need to be more complicated than those of non-adaptive networks. Thus a very complicated programme is required for the construction of these networks, which contains all the requisite information. An important and interesting question concerns the possibility of leaving the bulk of the selection process to be performed during the lifetime of the network, i.e., by adaptation. A. M. Uttley ([152], [153]) has outlined how certain replicated circuits might arise by chance in a randomly interconnected network, thus diminishing the amount of selective information which has to be supplied either "genetically" or "epigenetically". What remains is the problem of specifying the complicated functions required for functional coding. For high levels of functional coding, each formal neuron need only represent a random selection of the specified functions of the network, provided it correctly decodes its inputs. It is interesting that recent work on the design of adaptive machines capable of learning to classify, represent, and recognize patterns of stimuli, has given rise to machines whose functional organization is apparently very similar to the redundant networks we have designed. (see [149], [131], and [132]).
The combination of such theories of automata and information processing with what might be called machine theories of adaptation, provides a not unreasonable model of what might be the organization of those parts of brains concerned with

perception, learning, and perhaps with memory. We have cited Lashley's work on ablation effects. Lashley concluded that it was not possible to demonstrate the isolated localization of the memory trace or "engram" anywhere within the nervous system. He supposed that there was no special reservoir of cells which would serve as the seat of special memories, every instance of recall requiring the activity of millions of neurons. Moreover the same neurons which retained the engram must also participate in countless other activities. (see also [146], and [172]). Sperry has suggested a number of principles that might give rise to this property: multiple interconnexions between nerve-cells, the fidelity of the wirings being controlled by specific biochemical factors, much overlapping of interconnexions; multiple reinforcement of any given function from numerous different sources any one of which may itself be capable of sustaining the activity; reciprocal and surround inhibition between and among neurons; the arrangement of cortical circuits in vertical rather than horizontal dimensions and the bilateral duplication of the cerebral hemispheres. J. C. Eccles, [36] and A. E. Fessard, [42] have inferred from their own experimental studies that presynaptic and lateral inhibition between and among neurons are important features contributing to error-insensitivity. In considering changes in the "evoked potentials" in various areas of the brain, associated with the engramming of conditioned responses. E. R. John, [76], demonstrated an effect related to the replication of circuits and the delocalization of function which we have discussed, namely that evoked potentials following the engramming have similar shapes, even over many anatomically distinct regions of the brain.

In the light of these experiments we note certain logical requirements on the structure of the error-insensitive networks obtained by functional coding techniques. These are the existence of large numbers of both excitatory and inhibitory synapses at all units, the existence of many presynaptic "axo-axonal" interconnexions, a great deal of multiple interconnexion and overlap, and the existence of lateral inhibition between groups of neurons. Furthermore as a consequence of functional coding, there is an extensive representational system in these networks, any unit of which can be activated by many different patterns. Finally the reliable activation of any complete pattern of activity would require the synergic activity of many units. All these features of functionally coded networks appear to have experimental correlates, most of which are the result of experiments performed on the central nervous system prpeer. We should not expect to find high levels of functional coding in peripheral areas, but rather many replicated circuits more in line with the "telephone exchange" analogy. At some intermediate stage where the degree of synaptic interaction among neurons is sufficiently high to sustain functional coding as well as circuit replication there will emerge assemblies of synergic neurons, which may be taken to be the "functional units" of the network. Perhaps the vertically organized groups of neurons found in sensory projection areas ([108], [73]) may be taken to be the experimental correlates of these functional units.

We consider it to be of some importance that deductive models of the organization of neuronal networks be forthcoming which lead to the specification of such entities

as functional unit, engram, and so on. R. L. Gregory, [58], has made the point that a knowledge of function is required to classify observed biological features into "essential" or "accidental" properties, and that it is only when the functional units of the system being studied can be identified, that deductive inference and not mere description becomes possible. Gregory made the further point that the neurologist is never able to identify functional units directly by observing neurons and their interconnexions; in all cases knowledge is needed of what neurons do, and of how they do it. This raises a peculiar and difficult problem common to almost all attempts to apply cybernetics to biology. For in order to do this, one must know precisely what is the ensemble of possibilities upon which operates the selection process that alone gives meaning and utility to the ideas of message and information. For the neurological problem this is equivalent to saying that one needs to know the code or codes of the nervous system. Thus any real application of theories of formal neuronal networks can follow only from a knowledge of what neurons do and of how they do it. This takes us far from the ablation experiments of the physiological psychologist, to the electrophysiologists who measure not behaviour deficits and the like, but the firing patterns of neurons and the changing electrical potentials of nervous tissue, the ECoG and EEG. Once we enter the domain of experimental neurophysiology however, it becomes difficult to see how the automata models apply to data. In a recent review (J. D. Cowan, [28]) we have outlined the main results of the last twenty-five years work in neurophysiology. This review is included in this thesis as Appendix B. The reader who is unfamiliar with the subject matter is advised to read it before proceeding any further. For convenience we reproduce here several passages from the final discussion which summarize our viewpoint:

"... macroscopic studies of neuronal anatomy and physiology lead to the view of the nervous system as a complex of ordered nets arranged in a three level hierarchy... EEGs, ECoGs, and D.C. potential records indicate that there is an underlying dynamical activity... Feedback inhibition undoubtedly plays a very important role in maintaining the activity, which is controlled both by external and internal signals, especially from... reticular systems. The EEG serves as a measure of the general level of the dynamic(al) activity ,..., the D.C. potentials as measures of the stationary states of such activity..., and the ECoG and Evoked Responses as measures of ... incoming and outgoing fluxes (of excitation). There are two broad categories of evoked responses. One is associated with tonic aspects of the nervous system (and) is related to the ... spontaneous activity, and is probably generated in the same corticothalamic circuits. The other is the modality-specific response associated with phasic aspects ... (probably) generated in different circuits... Studies of isolated cells have established the manner in which cells respond to and emit spikes... there is (often) a monotonic relation between the rate of spike emission and the currents generated in somato-dendritic regions of the neuronal membrane. The

role of ... EPSPs and IPSPs ... in determining such responses is becoming known. There are definite relationships between spike patterns and macropotentials... It is probable that such potentials are the algebraic sums of EPSPs and IPSPs seen through (a)... low-pass filter, and that (the) D.C. potentials reflect the states of balance between EPSPs and IPSPs that constitute the stationary states of the dynamical system. Studies of the interactions between cells and of the phasic responses of cells to patterned stimuli have established that there is a specific organization of columnar nets maintained by nearest-neighbour interactions,(and by) spatial summation and inhibition. There are topological maps between many nets and from receptive fields. Studies of the effects of anaesthesia indicate that tonic responses may also be present in the cells of "physiological animals" although it is not clear that the precise columnar organization is then completely preserved. Concerning the neuronal coding of signals, one apparently general operating mode is (by)... pulse rate modulation, so that signals are deviations from (stationary) ... activity. The view that the Evoked Response signals by modulating the stationary balance between EPSPs and IPSPs comprising the D.C. potentials is consistent with this. However pulse rate modulation is only one possibility. Receptive field experiments indicate that the specificity of the individual neuronal response is a very important signal. These responses occur as spike bursts, and it is possible that in physiological preparations, variations in ... the frequency of bursts, (the) number of spikes per burst, cell specificity and response latency all constitute meaningful signals. One view is that almost all changes in the spatiotemporal patterns of firing within neuronal nets convey information. Per contra it has been suggested that nets average over incoming fluxes by processes of spatial and temporal summation and integration within cells, plus diffusion between cells, and that the shape of the Evoked Response ... is the important signal. Certainly receptive-field experiments indicate that signals (appearing) locally within nets are important for detecting and abstracting qualitative features of stimuli. Neither how these features are integrated into a holistic representation of stimuli, nor how such representations are eventually coded and stored, is yet known....

It is clear that the automata approach is seriously deficient in many respects. In the automata we have considered changes in the firing patterns occur only at instants specified by an external clock, i.e., they are already "synchronized" in the time domain. There is no clear evidence that the CNS operates in such a fashion. In any case the formal neuron is a rather crude abstraction of neuronal behaviour and the theory covers special aspects such as functional stability and doesn't serve to help us understand the responses of nets comprising many thousands or many millions of cells. The theory in fact was designed for the analysis of small-scale local interactions between abstract functional units, and was never intended for neurodynamics. Digital computer simulations of neuronal nets (see

[39], [144]) together with some combinatorial analysis ([144], [60]) have given us some indication of the type of activity to be found in homogeneous networks of randomly interconnected formal neurons. If there are only excitatory synapses in the net, the only stable states of activity are either a large proportion of all the units in the net are active, or else the net is quiescent. This is the "switching effect", discovered by many workers. It has been suggested that a neuronal network which acts itself as a switch might serve as a functional unit, and so correspond to one of McCulloch & Pitts' formal neurons. In case there are also inhibitory synapses, the behaviour is more complex and several intermediate stable levels of activity can persist. Farley has shown that even small networks comprising about one hundred cells can exhibit quite complicated behaviour, recruiting responses, augmenting responses, rhythms and so on, that are reminiscent of experimental phenomena. What is lacking in this approach however, are the concepts and mathematical tools that would further the analysis of the responses of very large networks, nor is there any real attempt to face the problems of coding in these networks. There has been one interesting attempt at a mathematical treatment of the responses of large nets, [12]. In this neurons are assumed to be randomly distributed in a mass with a given volume density. The neurons have thresholds, synaptic delays, EPSPs and IPSPs, and a summation time constant. Attention is directed to the proportion of cells becoming "sensitive" per unit time. Sensitive cells are those which are not refractory and can therefore be fired by a sufficiently potent stimulus. Although the mathematical treatment is not rigorous, the conclusions are essentially correct. The switching effect is discovered, and the conditions for the distortion-free propagation of plane waves of neuronal excitation are given. It is shown that the connexion function which gives such propagation is equivalent to that one found empirically by D. A. Sholl, [142]. Since only excitatory synapses are present, this wave propagation is unstable and depends critically on the stimulus, on the cell properties, and on the local density of interconnexions. Beurle introduced two mechanisms of interest. Waves were stabilised by "servo" control from external nets which acted by firing off cells ahead of the wavefronts. And cell thresholds were assumed to be dependent upon past activity. These extensions of the random network assumptions are important. The servo idea is an attempt to treat network interactions, the control of one network by another. Threshold modification by past activity changes the responses of a network, according to past responses. So the network can be trained, given the proper feedbacks, to act as a permanent store of messages. Indeed all the possibilities of behaviour, functional stability, adaptation, and so on, which we considered to be present in functionally stable automata, are to be found in these networks. However the scale has been changed, so to speak, in that complete networks play the role of functional units, and dynamical variables play an important role in the encoding of stimuli.
There is therefore the intriguing possibility that some kind of answers to the problems of modelling some aspects of the activity of neuronal networks in the CNS might be forthcoming from a mathematical formulation closely related to Beurles.

But there are a number of problems that must be solved before this approach can be made useful. The analysis has some defects. For example the effects of refractoriness, of the finite size of the network, and of delays, are not properly formulated. To some extent these defects are not too serious. Correct formulations of similar problems have been given by J. S. Griffith, [59], [61] and by M. Ten Hoopen, [70]. Their conclusions are similar to Beurle's concerning switching and stability, if somewhat less far-reaching in scope. However what is really lacking from the whole approach is that it does not make contact with the experimental variables of the neurophysiologist, the histograms, correlograms, ECoGs and so forth, and so in this sense it is not a falsifiable, or a fortiori, a testable model for the responses of neuronal networks in the central nervous system. Moreover while the image of interacting waves of neuronal excitation suggests many possibilities for the coding of messages, it has not been used so far to produce precise, quantitative predictions of how and where permanent changes take place in the networks of the central nervous system, and what they represent.
What is required then is a novel mathematical formulation of the responses of neuronal networks that takes account of many of the organizational features of the central nervous system that we have listed, in which the variables relate to measurable quantities by current experimental and data processing techniques, and in which there is a precise hypothesis concerning the nature of neuronal coding.

2. Formulation of the Neuronal Equation

In chapter 1 we have outlined theories of neuronal networks that were based on the McCulloch-Pitts neuron, and on the Beurle neuron. In this chapter we shall outline the construction of a mathematical theory that is aimed at a more exact description of the experimental phenomena, which can be correlated with the data. Our starting point is J. C. Eccles' well-known lumped equivalent circuit for the generation of post-synaptic potentials in the cell membranes of motoneurons ([35], [36]), and the less well-known but more accurate distributed equivalent circuit or "cable" equations for neuronal membrane of W. Rall, [120], [121]. Figure C.26 and Figure C.27 of Appendix B show the lumped equivalent circuit and its responses, and Figure C.28 of this Appendix shows the effects of distributed parameters upon the responses. The reader is referred to Appendix B for the details. Since we are interested in modelling the responses of large networks, we are concerned only with the crudest possible nontrivial way of writing down an equation for the responses of a single neuron. We therefore assume the model for a neuron shown in Figure C.1. The synapses are all "in parallel", and each delivers a current impulse roughly of the form shown in Figure C.2a to the dendritic cable. As a result of signal transmission through the cable, these current impulses are "smeared" so that when they finally impinge on the "threshold" membrane they have the form shown in Figure C.2b. (The exact nature of these impulses can be obtained from Rall's equations. Since these number four sets of simultaneous ordinary differential equations or an equivalent partial differential equation for each neuron, we are content to reason heuristically at this point, and to be guided mainly by physical intuition).

Let the current impulse shown in figure Figure C.2b be denoted by $i(t)$ where t is the epoch of time. Suppose that there are independent Shot-Noise inputs at each synapse. Then we can represent the mean current built up at the threshold membrane as

$$\langle I \rangle_T = \sum_s n_s \langle i \rangle_T \tag{2.1}$$

where we have assumed that the membrane is quiescent at the epoch 0 and the average is over the interval $[0, T]$. n_s is the number of inputs to the s-th synapse in the given interval. Equation 2.1 is of course the first part of Campbell's Theorem

(see W. Feller [41]). We may rewrite Equation 2.1 as:

$$\langle I \rangle_T = \sum_s v_s \int_0^T i(u)\,\mathrm{d}u\,, \tag{2.2}$$

where v_s is the mean "rate" at which inputs arrive at the s-th synapse, averaged over the interval of duration T units. (Note that we are already neglecting an important feature of neuronal responses. By assuming a common impulse response $i(t)$, we are neglecting changes in the time constants that occur both in spatial and in temporal summation which change the *latency* of the membrane response).
For the lumped model Equation 2.1 and Equation 2.2 reduce to:

$$\langle I \rangle_T = \sum_s (E)_s \cdot \delta_s \cdot g_s \cdot v_s \tag{2.3}$$

where $(E)_s$ is the battery driving the s-th synapse, δ_s is the operating time for the synapse, and g_s is the synaptic conductance. The batteries are well-known to be $-E_m$ at an excitatory synapse, and $E_I - E_m$ at an inhibitory synapse, where E_m is the resting potential of the membrane, and E_I is the equilibrium potential for an IPSP. Let us now assume that the input rates vary so that we can work with $\langle I \rangle_T$ as a function of the epoch t, rather than of the interval $[0, T]$. This means that they must be either rapidly changing or slowly varying with respect to the interval of duration T. We allow ourselves to choose the duration to satisfy this criterion. *And let us now approximate the effects of the electrotonic decrement represented in the cable equations, merely by an attenuation and a delay!* Thus we have another formula for the "current" built up at the threshold membrane:

$$I_r(t) = \sum_s \alpha_{sr} \cdot v_s\,(t - \tau_{sr}) \tag{2.4}$$

where $I_r(t)$ is the current built up in the r-th neuron, where τ_{sr} is the effective delay from the s-th to the r-th neuron (incorporating transit time, synaptic and dendritic delays), and where the coupling coefficients α_{sr} are defined as:

$$\alpha_{sr} = \begin{cases} -E_m \delta_e g_{sr} h_{sr}, & \text{excitatory synapse} \\ (E_I - E_m)\,\delta_i g_{sr} h_{sr}, & \text{inhibitory synapse.} \end{cases} \tag{2.5}$$

where h_{sr} is the attenuation factor for the sr-th interaction.
We now have to consider the way in which a response to such an excitation is elicited in the cell. The Hodgkin-Huxley equations ([69]) provide us with a way of calculating the response of axonal threshold membrane to voltage transients. Were we to assume that the above excitation acts directly on the axon-hillock of motoneurons, we could compute the cell's responses. This would entail the integration of several nonlinear differential equations. R. Fitzhugh, [44], has shown

how such equations give rise to threshold phenomena consistent with "all-or-none" responses. Appendix B, Figure C.29 and Figure C.30 show the equivalent circuit for the Hodgkin-Huxley equations, and the results of Fitzhugh's calculations.
In principle then, just as with the input excitation to a neuron, we could integrate another set of simultaneous nonlinear differential equations to obtain the distribution of impulses emitted by the neuron. However such is the complexity of the requisite computations (and it is by no means clear that the equations cover all the relevant phenomena) that *we have chosen to approximate the response by empirical curves obtained from experiments* on the mean responses of nerve membranes to stimulating currents. There are many experimental results. Figure C.31, Appendix B shows three typical curves ([68], [154], and [55]). The interpretation of the data is difficult since many different preparations were used, and effects such as adaptation and accommodation play a large role in determining the shape of the curves. However it often seems to be the case that over a wide range of current inputs, if the reciprocal of the first and second impulses elicited from a cell is taken as the mean firing rate, then the slope of the resulting rate-intensity curve changes in such a way that it can be reasonably fitted by a sigmoidal function. We therefore take as our fundamental formula for the mean rate $v(t)$ at which impulses are emitted by a neuron, the function:

$$v(t) = \frac{1}{r} \cdot \frac{1}{\left[1 + \exp\left(-\beta \frac{I(t)}{I_{\frac{1}{2}}(t)}\right)\right]} \tag{2.6}$$

where

$$\beta = \frac{I_{\frac{1}{2}}}{I_{\frac{1}{2}} - I_0} \cdot \ln\left(\frac{l - r \cdot v_0}{r \cdot v_0}\right) \tag{2.7}$$

Figure C.3 shows the resulting function. The constant r is the "refractory period" of the neuron, the threshold current $I_{\frac{1}{2}}$ is that current which drives the cell at the rate equal to $1/2r$ impulses per sec., the 50% level. I_0 is approximately $1.5I_{rh}$ where I_{rh} is the "rheobase", the smallest current which will elicit a single impulse from the neuron, and v_0 is the rate corresponding to the current I_0.
Two remarks are called for. Firstly there is no fundamental reason for choosing the curve of Equation 2.6, which is of course the well-known Logistic curve of demography ([111], [100]). It happens to be a convenient and tabulated sigmoidal function. What we are interested in studying is the qualitative nature of the results issuing from this choice of a sigmoidal function, not the exact quantitative aspects. There is no good reason for choosing any one function over another, and the Logistic curve is one of the simpler sigmoidal curves to work with. Secondly it will not have escaped the reader that by choosing the reciprocal of the interval between the first and second impulses emitted by a neuron, as a measure of the mean rate of impulse emission, we are ignoring adaptation and accommodation. In later work

we hope to show how these effects can be incorporated in the formulation. For the moment we must press on toward a manageable formulation of the equations of "motion" of neuronal networks.
Let us introduce the variable

$$X(t) = 1 - r \cdot v(t)\,. \tag{2.8}$$

Then we may combine Equation 2.4, Equation 2.6 and Equation 2.8 to give a difference equation relating the neuronal output to its inputs:

$$\boxed{\ln\left(\frac{X_r(t)}{1 - X_r(t)}\right) = \in_r + \frac{1}{\beta_r} \cdot \sum_s \alpha_{sr} \cdot X_s\left(t - \tau_{sr}\right)} \tag{2.9}$$

where

$$\in_r = \beta - \left(\frac{\beta}{r I_{r,1/2}}\right) \sum_s \alpha_{sr}$$

and

$$\frac{1}{\beta_r} = \frac{\beta}{r I_{r,1/2}} \tag{2.10}$$

The variable X can be interpreted as the fraction of time when a neuron that is continually emitting spikes is not refractory, i.e., the fraction of time when it is "sensitive" to incoming stimuli. For this reason we call X the *sensitivity* of the neuron. Clearly the variable $(1 - X)$ represents the fraction of time when the neuron is "refractory" and cannot be fired by incoming stimuli, i.e. the insensitivity. For sufficiently large samples we can use a central result of renewal theory, [29], that the renewal density, or here the mean interspike interval τ is inversely proportional to the mean rate v, so that we can also write X as

$$X = 1 - \frac{r}{\tau(t)} \tag{2.8'}$$

thus relating sensitivity to the interval between spikes. The sensitivity function is in fact related to certain functions occurring in the renewal theory of Geiger counters. A Type I counter is one which is blocked or refractory following the reception of a particle, for a certain time. The neuron may be thought of as the combination of a Type I counter with a coincidence detector. Let r be the refractory or blocked time for the system and v the rate of arrival of particles (or more correctly the rate of arrival of θ particles where θ is the threshold of the detector). It is well-known (see [29]) that the mean "failure time" when the system is insensitive is $r + 1/v$ and that N_t, the number of coincidences counted in time t, is asymptotically normally distributed with mean $v \cdot t/(l + rv)$. (This may not be quite the correct way of combining coincidence detectors with Type I counters but it seems reasonable). A large-sample estimate for v can be obtained by equating $v \cdot t/(l + r \cdot v)$ with N_t.

Thus we have:

$$\frac{v \cdot t}{1 + r \cdot v} = N_t \tag{2.11}$$

or,

$$v = \frac{N_t/t}{1 - r \cdot N_t/t} \,. \tag{2.12}$$

If we identify the counting rate N_t/t of the system with $v(t)$, the rate of emission of spikes from a neuron, we see that the estimate of the rate of arrival of sufficiently many inputs to fire the neuron is given by the quantity:

$$\frac{1}{r} \cdot \frac{1 - X}{X} \,.$$

This provides a heuristic justification for using the sensitivity function, although it is certainly not a rigorous way of connecting X with renewal theory. More work is required here.
We now wish to convert the fundamental difference Equation 2.9 into a differential equation. The motivation for this will be made explicit in the next two chapters. We make the assumption that *all the delays are the same*, so that we have $\tau_{sr} = \tau_r$, and we rewrite Equation 2.9 as:

$$\ln\left(\frac{X_r\left(t + \tau_r\right)}{1 - X_r\left(t + \tau_r\right)}\right) = \in_r + \frac{1}{\beta_r} \cdot \sum_s \alpha_{sr} \cdot X_s(t) \,. \tag{2.13}$$

Since the function $\ln\left(\frac{x}{1-x}\right)$ is a continuous well-behaved function we can expand it in a Taylor series and for sufficiently small τ_r we have the equation:

$$\ln\left(\frac{X_r(t)}{1 - X_r(t)}\right) + \tau_r \cdot \frac{\mathrm{d}}{\mathrm{d}t}\left\{\ln\left(\frac{X_r(t)}{1 - X_r(t)}\right)\right\} = \in_r + \frac{1}{\beta_r} \cdot \sum_s \alpha_{sr} \cdot X_s(t)$$

or,

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}T}\left\{\ln\left(\frac{X_r(T)}{1 - X_r(T)}\right)\right\} + \ln\left(\frac{X_r(T)}{1 - X_r(T)}\right) = \in_r + \frac{1}{\beta_r} \cdot \sum_s \alpha_{sr} \cdot X_s(T)} \tag{2.14}$$

where

$$T = \frac{t}{\tau_r} \,.$$

This then is our neuronic equation, an ordinary nonlinear differential equation in the variable X. In the next chapter we shall outline some of its properties. But it is clear that our neuronal model differs considerably from the formal neuron described in chapter 1. The formal neuron is of course a model of the "black-box" variety, serving to model "input-output" patterns of neuronal activity. We have chosen to model some of the details inside the black-box, which have been

discovered by electrophysiology. To do justice to such details would require not one neuronic equation, but as we have indicated, many simultaneous equations for each neuron. Our approximate equation serves the purpose of modelling networks, and its validity is to be tested by the accuracy of its predictions concerning the responses of neurons within networks, and by the conceptual insights it may provide about the qualitative features of the activity of these networks.

3. Neuronal Oscillations

We now start along the classical pathway followed by every theoretical physicist who wishes to study complicated dynamical systems, we set out to construct harmonic oscillators! Let us therefore examine the function $\ln\left(\frac{x}{1-x}\right)$ as a function of x, for $0 < x < 1$. (see Figure C.4). It will be seen that there are effectively three regions where the linear approximation is not unreasonable: these are the intervals (0.005, 0.015), (0.3, 0.7), and (0.985, 0.995). Thus there are two very narrow regions and one fairly broad region within which we can approximate to $\ln\left(\frac{x}{1-x}\right)$. For the middle region we use the series expansion, see [1]:

$$\ln z = 2 \cdot \left\{ \frac{z-1}{z+1} + \frac{1}{3}\left(\frac{z-1}{z+1}\right)^3 + \frac{1}{5}\left(\frac{z-1}{z+1}\right)^5 + \ldots \right\} \tag{3.1}$$
$$\mathscr{R}(z) \geqslant 0,\ z > 0\,,$$

which becomes

$$\ln\left(\frac{x}{1-x}\right) = 2 \cdot \left\{ (2x-1) + \frac{1}{3}(2x-1)^3 + \frac{1}{5}(2x-1)^5 + \ldots \right\} \tag{3.2}$$
$$\mathscr{R}\left(\frac{x}{1-x}\right) \geqslant 0,\ \frac{x}{1-x} > 0\,.$$

or in the linear approximation

$$\ln\left(\frac{x}{1-x}\right) = 4x - 2\,, \quad x \in (0.3, 0.7)\,. \tag{3.3}$$

By graphical means we have fitted to the other regions the lines:

$$\ln\left(\frac{x}{1-x}\right) = \begin{cases} 139 \cdot x - 6, & x \in (0.005, 0.015) \\ 139 \cdot x - 133, & x \in (0.985, 0.995). \end{cases} \tag{3.4}$$

As we shall indicate the consequences of this linearization are not trivial. For these three cases then, Equation 2.14 can be rewritten as:

$$m_r \cdot \frac{\mathrm{d}X_r}{\mathrm{d}T} + m_r X_r + c_r = \in_r + \frac{1}{\beta_r} \cdot \sum_s \alpha_{sr} \cdot X_s\,. \tag{3.5}$$

We consider first the elementary case of two interacting neurons. We suppose that one neuron is excitatory $\alpha_{21} = -E_m \delta_e g_{21} h_{21}$, and that the other is inhibitory, $\alpha_{12} = (E_I - E_m)\, \delta_i g_{12} h_{12}$. We now have two equations:

$$m_1 \cdot \frac{\mathrm{d}X_1}{\mathrm{d}T} + m_1 X_1 + c_1 = \in_1 + \frac{1}{\beta_1} \cdot \alpha_{21} X_2 \tag{3.6a}$$

$$m_2 \cdot \frac{\mathrm{d}X_2}{\mathrm{d}T} + m_2 X_2 + c_2 = \in_2 + \frac{1}{\beta_2} \cdot \alpha_{12} X_1 \tag{3.6b}$$

Let us first look at the stationary states of this system, i.e., the states $X_i = q_i$, such that $X_i = 0$. These are the solutions of the equations:

$$m_1 q_1 + c_1 = \in_1 + \frac{1}{\beta_1} \cdot \alpha_{21} q_2 \tag{3.7a}$$

$$m_2 q_2 + c_2 = \in_2 + \frac{1}{\beta_2} \cdot \alpha_{12} q_1 \,. \tag{3.7b}$$

We recall that $\in_r = \beta - \frac{1}{\beta_r} \cdot \sum_s \alpha_{sr}$. To guarantee that there exist stationary states in the requisite ranges of X, we need an additional degree of freedom in the coefficients $\in_r$. So *we introduce the external control signals* K_r for this purpose, so that now

$$\in_r = \beta - \frac{1}{\beta_r} \cdot \sum_s \alpha_{sr} + K_r \tag{3.8}$$

and by suitable choice of K_r we can obtain the requisite stationary states. The reader will have noted that what we have constructed is a kind of *negative feedback circuit* with a nonlinear input that controls the level of activity in the circuit. Figure C.5 shows such a circuit. We can combine equations Equation 3.6a, Equation 3.6b, Equation 3.7a, Equation 3.7b and Equation 3.8 to give:

$$\beta_1 m_1 \cdot \frac{\mathrm{d}Y_1}{\mathrm{d}T} + \beta_1 m_1 Y_1 = \alpha_{21} Y_2 \tag{3.9a}$$

$$\beta_2 m_2 \cdot \frac{\mathrm{d}Y_2}{\mathrm{d}T} + \beta_2 m_2 Y_2 = \alpha_{12} Y_1 \tag{3.9b}$$

where

$$Y_r = X_r - q_r \,. \tag{3.10}$$

We recognize these equations as defining a damped harmonic oscillator. (Recall that α_{21} is positive and α_{12} negative). Instead of proceeding in the standard way however, we shall introduce the Hamiltonian formalism, in preparation for what is to follow.
Let us introduce the variables:

$$Z_r = Y_r \cdot e^T \tag{3.11}$$

then it is obvious that

$$\frac{\mathrm{d}Z_r}{\mathrm{d}T} = \left(\frac{\mathrm{d}Y_r}{\mathrm{d}T} + Y_r\right) \cdot e^T$$

so that Equation 3.9a and Equation 3.9b can be transformed into:

$$\beta_1 m_1 \cdot \frac{\mathrm{d}Z_1}{\mathrm{d}T} = \alpha_{21} Z_2 \tag{3.12a}$$

$$\beta_2 m_2 \cdot \frac{\mathrm{d}Z_2}{\mathrm{d}T} = \alpha_{12} Z_1 \tag{3.12b}$$

which we recognize as the equations for undamped harmonic oscillations in Z_r. If we add the condition that

$$\alpha_{12} + \alpha_{21} = 0 \tag{3.13}$$

we can re-express Equation 3.12a and Equation 3.12b in the form:

$$\frac{\alpha_{12}}{2} \cdot \left(\beta_1 m_1 Z_1^2 + \beta_2 m_2 Z_2^2\right) = G \tag{3.14}$$

where G is a *constant of the motion* of the system. If we now transform back to the original X_r variables, we see that

$$G = \frac{\alpha_{12}}{2} \cdot \left(\beta_1 m_1 \cdot (X_1 - q_1)^2 + \beta_2 m_2 \cdot (X_2 - q_2)^2\right) \cdot e^T$$

is an explicit function of the time $T = t/\tau$. This means that G is not the *energy-function* for the system (see [50]) and indeed that the oscillations in X_r about the stationary states q_r are damped with a time constant τ.
In this simple case of course we can solve Equation 3.9a, Equation 3.9b or Equation 3.12a, Equation 3.12b directly to give the solutions

$$Z_r(T) = A \cdot \cos\left(BT + \varphi\right) \tag{3.15a}$$

$$X_r(T) - q_r = Ae^{-T} \cdot \cos\left(BT + \varphi\right) \tag{3.15b}$$

$$X_r(T) - q_r = Ae^{-t/\tau} \cdot \cos\left(BT/\tau + \varphi\right) \tag{3.15c}$$

where

$$B^2 = \frac{\alpha_{12}^2}{\beta_1 \beta_2 m_1 m_2}$$

so that the period of the oscillations is simply:

$$\lambda = 2\pi\tau/B = \frac{2\pi\tau\sqrt{\beta_1 \beta_2 m_1 m_2}}{\alpha_{12}} \tag{3.16}$$

and we have to be careful to ensure that $X_r(0)$ is in the correct range.
But what have we obtained with this machinery? We have constructed a linear damped oscillator whose damping constant is of the order of a few milliseconds at

most. This follows directly from the fact that the fundamental neuronic equation is a difference equation. There is no inertia in the system and therefore no intrinsic mechanism for storing "energy" in the oscillations. How are we to construct an undamped oscillator from such unpromising units? We do so as follows. *We introduce auxiliary feedback loops for each individual neuron.*
Consider the linearized version of the difference Equation 2.13, which takes the form:

$$m_r \cdot X_r\left(t+\tau_r\right)+c_r=\in_r+\frac{1}{\beta_r}\cdot\sum_s \alpha_{sr}\cdot X_s(t)\,. \tag{3.17}$$

Suppose that there exists a subset of "interneurons" indexed by s' that serve only to re-excite the r-th neuron. Figure C.6 shows what is meant. These interneurons are represented by the equations:

$$m_{s'} \cdot X_{s'}\left(t+\tau_{s'}\right)+c_{s'}=\in_{s'}+\frac{1}{\beta_{s'}}\alpha_{rs'}\cdot X_r(t)\,. \tag{3.18}$$

We can combine these equations with Equation 2.13 to give:

$$m_r \cdot X_r\left(t+\tau_r\right)+c_r=\in_r+\frac{1}{\beta_r}\cdot\left(\sum_s \alpha_{sr}X_s(t)+\sum_{s'}\alpha_{s'r}X_{s'}(t)\right)$$

which gives

$$\begin{aligned} m_r \cdot X_r\left(t+\tau_r\right)+c_r=\ &\in_r+\frac{1}{\beta_r}\cdot\sum_s \alpha_{sr}X_s(t)+\frac{1}{\beta_r}\cdot\sum_{s'}\alpha_{s'r}\frac{1}{m_{s'}} \\ &\cdot\left(\left(\in_{s'}-c_{s'}\right)+\frac{1}{\beta_{s'}}\alpha_{rs'}\cdot X_r\left(t-\tau_{s'}\right)\right) \end{aligned} \tag{3.19}$$

This equation may be rewritten as:

$$\begin{aligned} &X_r\left(t+\tau_r\right)-\frac{1}{m_r\beta_r}\cdot\sum_{s'}\left(\frac{\alpha_{s'r}\alpha_{rs'}}{m_{s'}\beta_{s'}}\right)\cdot X_r\left(t-\tau_s\right) \\ &=\left(\frac{\in_r-m_r}{m_r}+\frac{1}{m_r\beta_r}\sum_{s'}\frac{\alpha_{s'r}}{m_{s'}}\cdot\left(\in_{s'}-c_{s'}\right)\right)+\frac{1}{m_r\beta_r}\cdot\sum_s\alpha_{sr}\cdot X_s(t)\,. \end{aligned} \tag{3.20}$$

We consider the left-hand side of this equation and expand the retarded and advanced X_r about the epoch t, so that for sufficiently small delays, τ_r and τ_s, we can write this as:

$$X_r(t)=\tau_r\cdot\frac{\mathrm{d}X_r}{\mathrm{d}t}-\frac{1}{m_r\beta_r}\cdot\sum_{s'}\left(\frac{\alpha_{s'r}\alpha_{rs'}}{m_{s'}\beta_{s'}}\right)\cdot\left(X_r(t)-\tau_{s'}\cdot\frac{\mathrm{d}X_r}{\mathrm{d}t}\right)$$

or as:

$$\left[\tau_r + \frac{1}{m_r\beta_r}\cdot\sum_{s'}\tau_{s'}\cdot\left(\frac{\alpha_{s'r}\alpha_{rs'}}{m_{s'}\beta_{s'}}\right)\right]\frac{\mathrm{d}X_r}{\mathrm{d}t} + \left[1 - \frac{1}{m_r\beta_r}\cdot\sum_{s'}\left(\frac{\alpha_{s'r}\alpha_{rs'}}{m_{s'}\beta_{s'}}\right)\right]X_r(t)\,.$$

It is now obvious that there will be an energy storage and no damping whenever

$$\sum_{s'}\left(\frac{\alpha_{s'r}\alpha_{rs'}}{m_{s'}\beta_{s'}}\right) = m_r\beta_r\,. \tag{3.21}$$

Since we are free to pick the coupling coefficients in whatever way we please it is clear that solutions to this equation can be found. It is also obvious that there are many more possibilities using not one interneuron per loop but many, and the equations will be considerably more complicated, involving higher derivatives of X_r. We shall here consider only the simplest case of one interneuron per loop and all delays τ_s equal, in which case equations Equation 3.20 can now be written as:

$$\boxed{\frac{\mathrm{d}X_r}{\mathrm{d}T} = \gamma_r + \frac{1}{m_r\beta_r}\cdot\sum_s \alpha_{sr}X_s} \tag{3.22}$$

where

$$T = \frac{t}{\tau_r + \tau_s}$$

and where

$$\gamma_r = \frac{\in_r - m_r}{m_r} + \frac{1}{m_r\beta_r}\cdot\sum_{s'}\left(\frac{\alpha_{s'r}}{m_{s'}}\right)\cdot(\in_{s'} - c_{s'})\ . \tag{3.23}$$

Equation 3.22 is then *the fundamental differential equation for neuronal responses.* It should be remembered however that this equation is only valid for the ranges of X_r already listed, in the sense that only in these specified ranges does the linear approximation to the logistic curve hold. It is tempting to assume that we can in fact write down the full nonlinear differential equation:

$$\frac{\mathrm{d}}{\mathrm{d}T}\left\{\ln\left(\frac{X_r}{1-X_r}\right)\right\} = \gamma_r + \frac{1}{\beta_r}\cdot\sum_s \alpha_{sr}X_s \tag{3.24}$$

valid for these same ranges. This is of course approximately true in the linear approximation. But if we try to follow the same course of using the interneurons to balance the nonlinear damping inherent in equations Equation 2.14, we find that there is an important difference between the linear and the nonlinear cases. As we have seen, the energy balance is contained in equation Equation 3.21. Thus *the solution is independent of* X_r. A little reflection will show that this cannot be the case for the nonlinear equations and that a transcendental equation has to be solved which cannot have solutions at more than a few isolated points. Although

we have carried out these calculations we do not include them here, the main reason being that we are not really interested in solutions which depend critically upon X, for we are eventually going to immerse our circuits in a "heat-bath" of many other circuits and signals, and we are going to consider only the ensemble and time averages that issue from such an immersion.
In what follows we shall however consider both the linear and the nonlinear differential equations in the specified regions, in order to keep a check on the validity of the linearized calculations. Therefore we shall proceed with some development of Equation 3.24 obtaining the linearized solutions by specialization, and show how the equations are used to model the activity of neuronal networks in the central nervous system.

4. Corticothalamic Networks

Before plunging any further into mathematical pyrotechnics it is necessary to say a little concerning what we are attempting to model. Although this thesis is primarily concerned with an exposition of the theoretical notions involved in modelling brain activity, it is well to state that these notions are closely tied to experimental phenomena, and that the theory is supposed to be descriptive. The reader who is familiar with the contents of Appendix B does not need to be reminded of the importance attached to the circuits between cortical and thalamic networks. Such circuits are held to generate most of the spontaneous activity of the central nervous system, to be involved in scanning, perception, learning, and all the rest, but to this date there has been no real treatment of corticothalamic networks in the theoretical literature, although D. O. Hebb, [67], gave a very plausible, if somewhat heuristic theory of behaviour, based on certain "cell assemblies" which were in fact corticothalamic assemblies. In what follows we shall use equations Equation 3.22 and Equation 3.23 to model the "tonic" activity supposedly generated in these circuits (see Appendix B for details). Figure C.7 shows what is meant. We suppose thalamic neurons to act on the apical dendrites of pyramidal neurons in the cortex, via what are known as the "non-specific afferents". The post-synaptic potentials induced by afferent stimulation, all assumed to be EPSPs, supposedly travel down the dendrites to the cell body, and thence summate to fire the neuron. These neurons in turn synapse via corticifugal fibres, back in the thalamus, in a "non-specific" nucleus, thus completing a closed corticothalamic circuit. In addition both the cortical pyramidal cells and the thalamic counterparts are imbedded in a plexus of interneurons, that serves the purpose of acting as an "energy" store. Figure C.8 demonstrates that there is indeed such a plexus in the cerebral cortex. Finally there is constant bombardment of both sets of neurons from other networks, and there are control signals and Stimuli all playing on the elements comprising the network. What we are about to do is to investigate the consequences of this "vertical" circuitry (see [146]), in the belief that the vertical organization is of the utmost importance in the organization of the central nervous system. But how are we to proceed? With such a complicated system in mind we cannot hope to analyse in detail any of the exact features of the activity. Indeed the solution for the case of a single neuronal oscillator will not tell us anything very useful about the behaviour of any sizeable population of interacting oscillators. (This of course is what we are now aiming at, a very large number of interacting circuits, although we do not specify how many circuits we are considering, just that there are many). The reason of course is that the individual oscillator rapidly loses its identity as the

size of the network increases. In a network with many components, it is no longer possible to specify the "state" of the network ($X_r = a_r,\ r = 1, 2, \ldots, N, \ldots, 2N$) at any epoch T, simply because the large network is neither completely observable nor completely controllable. Even in the unlikely event that we have a reasonable solution for the behaviour of the network, we should be unable to use it, precisely because of the inherent indeterminacy of the "initial state" of such a network. In certain cases however, time-averages of quantities defined on the network provide meaningful statistics relating to the overall or "macroscopic" behaviour of the network. *Statistical mechanics* (see J. W. Gibbs [49]) provides us with one method for calculating these averages, provided that certain requirements are satisfied. The method is well-known: the single undamped oscillator evolving in a unique way from a precisely determined initial state must be replaced by the idea of a set of oscillators, distributed according to some stochastic law, each evolving in a deterministic manner. We must consider therefore, in place of the single oscillator, an *ensemble* of oscillators, and we must be satisfied with the information which may be obtained from the study of such an ensemble, namely the moments of probability distributions associated with the ensemble. Provided these distributions correspond to stochastic processes that are stationary and ergodic, we may equate ensemble averages with time averages, and the job is done.
It will be seen however, that before we can proceed to link the mechanical with the stochastic aspects of the problem, we must be sure that we are dealing with undamped oscillators. This is why we were so concerned in chapter 3 to find ways of eliminating damping. In what follows we shall treat Equation 3.24 as if they were valid for all the range of X_r, after which we shall investigate what modifications have to be introduced to eliminate those ranges of X from the solutions, in which we know that damping does occur. Consider therefore Equation 3.24 in the form:

$$\frac{\mathrm{d}X_r}{\mathrm{d}T} = \left(\gamma_r + \frac{1}{\beta_r} \cdot \sum_s \alpha_{sr} X_s\right) X_r \cdot (1 - X_r) \tag{4.1}$$

subject to the auxiliary conditions Equation 3.8 and Equation 3.23, with slight modifications in the format.
We now consider a system of N excitatory neurons (thalamic cells) interacting with a system of N inhibitory neurons (cortical cells), together with all the associated interneurons and control inputs. (see Figure C.8). The equations for such a network are those of Equation 4.1. These equations can be transformed in an interesting way. Let us introduce the new variable:

$$u_r = \ln\left(\frac{\frac{X_r}{q_r}}{1 - X_r}\right) . \tag{4.2}$$

Then Equation 4.1 becomes:

$$\beta_r \cdot \frac{\mathrm{d}u_r}{\mathrm{d}T} = \sum_s \alpha_{sr} \cdot \left(\frac{q_s \cdot e^{u_s}}{1 + q_s \cdot e^{u_s}} - q_s \right) \tag{4.3}$$

or

$$\frac{\mathrm{d}u_r}{\mathrm{d}T} = \sum_s \cdot \left(\frac{\alpha_{sr}}{\beta_s \beta_r} \right) \cdot \frac{\partial}{\partial u_s} \left(\sum_r \beta_r \cdot \left[\ln\left(1 + q_r \cdot e^{u_r}\right) - q_r u_r \right] \right)$$

or

$$\frac{\mathrm{d}u_r}{\mathrm{d}T} = \sum_s \left(\frac{\alpha_{sr}}{\beta_s \beta_r} \right) \cdot \frac{\partial \mathscr{H}}{\partial u_s} \tag{4.4}$$

where

$$\mathscr{H} = \sum_r \beta_r \cdot \left[\ln\left(1 + q_r \cdot e^{u_r}\right) - q_r u_r \right] . \tag{4.5}$$

$\mathscr{H}$ is of course the "Hamiltonian" of the network, is a constant of motion, and since it is not an explicit function of T plays the role of "energy" in the system. Thus we have constructed a conservative dynamical system that exhibits undamped oscillations in the variables X_r (or the "canonical" variables u_r) about the stationary states q_r. These states are either (trivially) 0 or 1, or else they are given by the equations:

$$\beta_r \gamma_r + \sum_s \alpha_{sr} q_s = 0 \tag{4.6a}$$

i.e.

$$q_s = \sum_r A'_{rs} \gamma_r \beta_r \,, \tag{4.6b}$$

where A'_{rs} is the rs-th element of the inverse of α_{rs}. By assumption this inverse exists since α_{sr} is skew-symmetric and of even order. Suitable choice of control currents and parameters ensures that such stationary states fall within the open interval (0,1).
The skew-symmetry follows of course from the assumption that the cortical cells are inhibitory and the thalamic ones excitatory ($\alpha_{rs} + \alpha_{sr} = 0$). It is easily shown that $\mathscr{H}$ is a positive-definite function of the canonical variables u_r, so that it satisfies all the formal requirements of Hamiltonian mechanics (see [50]). So we have obtained in the special case considered, an equation of motion for the activity of a finite population of interacting neurons that is essentially in Hamiltonian form, representing a system of undamped oscillators. The oscillations are in the sensitivity of the neurons or effectively in the interspike intervals.

In passing it is not without interest that N. Wiener and W. H. Pitts in the late 1940's appear to have been working along similar lines, but apart from the remarks quoted below nothing was published by them on the topic. All we have is the following remark of Wiener's:

The anatomical picture of the cortex suggests that we may usefully employ statistical methods in the study of its function. This work has been taken up brilliantly by Walter Pitts. He finds that under many conditions, the approximation to the activity of the cortex should not be made on a basis of rest and the zero basal activity. Proceeding from this he has developed the wave equations of the cortex. These promise to give an interpretation of the electrocorticogram. He has also suggested the exploration of the connexions between different parts of the cortex by suitable stimulation of regions and the observance of the frequency response at several points.

(N. Wiener [166])

Wiener's remarks were actually made at a 1946 conference. In my view it is remarkable that in the twenty years since, with the exception of Wiener's own rather abstract work on Hamiltonians and neuromechanics, [168], there is no trace of Hamiltonians and Statistical Mechanics in the field of neuronal modelling.

5. Statistical Mechanics of Systems of Neuronal Oscillators

We are now ready to forge the link between the mechanical and the stochastic aspects of the mathematical problem. This link is forged through the use of the "phase space" of the "incompressible" canonical variables $u_r (r = 1, 2, \ldots, 2N)$. In such a phase space the ensemble behaves like an incompressible fluid, so that the function $\int\int\ldots\int \mathrm{d}u_1\,\mathrm{d}u_2\ldots\mathrm{d}u_{2N}$ is an *integral invariant*. This invariant is a volume element of the space, so that *any* function of the form $\int\int\ldots\int D\,\mathrm{d}u_1\,\mathrm{d}u_2\ldots\mathrm{d}u_{2N}$ is also an integral invariant whenever D is an arbitrary function of the $2N$ constants of motion of the equations. It is well-known that Hamiltonian systems admit such invariants, [164].
The equation determining these invariants takes the form:

$$\frac{\partial D}{\partial T} + D \cdot \sum_r \frac{\partial X_r}{\partial u_r} = 0 \tag{5.1}$$

for differential equations of the form

$$\dot{u}_r = Y_r \quad (r = 1, 2, \ldots, 2N)\,. \tag{5.2}$$

It is obvious that the transformations Equation 5.2 reduce Equation 5.1 to that form of Equation 4.4 wherein $\frac{\partial Y_r}{\partial u_r} = 0$ for all r. Thus any function D of the constants of motion of Equation 5.1 will provide us with an integral invariant. It can then be shown that $\frac{\partial D}{\partial T} = 0$, i.e., that the ensemble is stationary as required. The ensemble average of any function $f(u_1, u_2, \ldots, u_{2N})$ is then defined as:

$$\bar{f} = \frac{\int\int\ldots\int Df\,\mathrm{d}u_1\,\mathrm{d}u_2\ldots\mathrm{d}u_{2N}}{\int\int\ldots\int D\,\mathrm{d}u_1\,\mathrm{d}u_2\ldots\mathrm{d}u_{2N}}\,. \tag{5.3}$$

As for ergodicity, it was shown by G. D. Birkhoff and others ([13]), that the great majority of non-integrable Hamiltonian systems (including ours!) possess what is sufficient for our purposes, the "quasi-ergodic" property. We are thus in a position to compute the time averages $\langle t \rangle$ of the variables associated without oscillators, provided we pick a suitable ensemble.

Let us therefore introduce as our invariant, the function:

$$\int \int \dots \int \rho \, \mathrm{d}u_1 \, \mathrm{d}u_2 \dots \mathrm{d}u_{2N}$$

where ρ is the *density* of the members of the ensemble within the volume element $\mathrm{d}u_1 \, \mathrm{d}u_2 \dots \mathrm{d}u_{2N}$. If this integration is taken over all phase spaces, we obtain the total number of members in the ensemble, say n. We shall work with the *relative density*, or density of the ensemble $\rho' = \rho/n$, so that our ensemble averages become

$$\bar{f} = \int f \rho' \, \mathrm{d}\tau, \quad \int \rho' \, \mathrm{d}\tau = 1, \quad \mathrm{d}\tau = \mathrm{d}u_1 \, \mathrm{d}u_2 \dots \mathrm{d}u_{2N} \, . \tag{5.4}$$

We have evidently constructed an invariant probability measure (see [89]) with ρ' playing the role of a probability density function. We shall now take ρ' to be a function of the Hamiltonian $\mathscr{H}$, although of course this is not the most general way of obtaining an invariant measure. Thus we have:

$$\rho' = \rho'(\mathscr{H}), \quad \mathscr{H} = \sum_r \beta_r \cdot \left[\ln \left(1 + q_r \cdot e^{u_r} \right) - q_r u_r \right] \, . \tag{5.5}$$

Let us now assume that we are dealing with a system comprising of a very large number of members, that the system is isolated and in statistical equilibrium, and let us fix our attention on a subset of the system. It is well-known that for such subsets, the probability density takes the form (see [62])

$$\rho' = e^{\frac{\chi - \mathscr{H}}{\theta}} \tag{5.6}$$

where χ and θ are constants such that

$$e^{-\frac{\chi}{\theta}} = \int e^{-\frac{\mathscr{H}}{\theta}} \, \mathrm{d}\tau \, .$$

The constant of motion of the whole system is $\mathscr{H} + \mathscr{H}'$ (subset plus complement). This distribution is of course the familiar "*canonical distribution*" of statistical physics (see [49]).
In what is to follow then, we shall be investigating the average (or most probable) behaviour of a network of interacting neurons supposedly part of a very large assembly of neuronal networks, whose dynamics as a result of feedbacks in the network, consists of undamped oscillations about certain stationary values of the mean rates at which neurons emit impulses. Moreover a certain constant function of the logarithms of these rates exists, and the probability of finding the network as a whole in any given state, as measured by time averages of certain functions yet to be defined, is itself a function of this constant.

5.1. Calculation of "canonical" averages

We now use the standard machinery that we have set up to compute ensemble averages with respect to the canonical distribution.

(a) The average of the function $A_r = \frac{\partial \mathscr{H}}{\partial u_r}$:

$$\begin{aligned}\bar{A}_r =& \frac{\int \frac{\partial \mathscr{H}}{\partial u_r} \cdot e^{-\frac{\mathscr{H}}{\theta}} \, \mathrm{d}\tau}{\int e^{-\frac{\mathscr{H}}{\theta}} \, \mathrm{d}\tau} \\ =& \frac{-\theta \int \frac{\partial}{\partial u_r} \left(e^{-\frac{\mathscr{H}_r}{\theta}} \right) \mathrm{d}u_r}{\int e^{-\frac{\mathscr{H}_r}{\theta}} \, \mathrm{d}u_r} \\ =& \frac{-\theta \left[e^{-\frac{\mathscr{H}_r}{\theta}} \right]_{u_r}}{\int e^{-\frac{\mathscr{H}_r}{\theta}} \, \mathrm{d}u_r}\end{aligned}$$

Since $u_r = \ln\left((X_r/q_r) \,/\, (1 - X_r) \right)$ according to Equation 4.2 and X_r runs from 0 to 1, u_r runs from $-\infty$ to $+\infty$. Hence

$$\begin{aligned}\left[e^{-\frac{\mathscr{H}_r}{\theta}} \right]_{u_r} &= \left[e^{-\frac{\beta_r}{\theta} \cdot \ln(1 + q_r \cdot e^{u_r}) - q_r u_r} \right]_{u_r} \\ &= \left[\left(\frac{1}{e^{-q_r u_r} + q_r \cdot e^{(1-q_r)u_r}} \right)^{\frac{\beta_r}{\theta}} \right]_{-\infty}^{+\infty} \\ &= 0 \,.\end{aligned}$$

Thus

$$\bar{A}_r = 0 \,.$$

But

$$\begin{aligned}A_r = \frac{\partial \mathscr{H}}{\partial u_r} &= \beta_r q_r \cdot \left(\frac{e^{u_r}}{1 + q_r \cdot e^{u_r}} - 1 \right) \\ &= \beta_r \cdot (X_r - q_r) \,.\end{aligned}$$

Thus we have the result that *the ensemble average of X_r equals the stationary state q_r.*
Let us also compute directly the time-average of X_r. We have:

$$\langle X_r \rangle = \lim_{T \to \infty} \left(\frac{1}{T} \cdot \int_0^T X_r(u) \, \mathrm{d}u \right) \,.$$

It follows immediately from Equation 4.1 that

$$\frac{\beta_r}{T}\ln\left(\frac{Y_r(t)}{Y_r(0)}\right) = \gamma_r\beta_r + \sum_s \alpha_{sr}\cdot\left(\frac{1}{T}\cdot\int_0^T X_s(u)\,\mathrm{d}u\right)$$

where

$$Y_r = \ln\left(\frac{X_r}{1-X_r}\right).$$

Thus

$$\beta_r\gamma_r + \sum_s \alpha_{sr}\langle X_s\rangle = 0\,. \tag{5.7}$$

Comparing this with equations Equation 4.6b determining the q_s, we see that:

$$\langle X_r\rangle = q_r$$

and hence we have the result that *the time-average of X_r equals the canonical ensemble average.*

(b) The average of the function $B_r = u_r\cdot\frac{\partial\mathscr{H}}{\partial u_r}$:

$$\begin{aligned}\bar{B}_r &= \frac{\int u_r\frac{\partial\mathscr{H}}{\partial u_r}\cdot e^{-\frac{\mathscr{H}}{\theta}}\,\mathrm{d}\tau}{\int e^{-\frac{\mathscr{H}}{\theta}}\,\mathrm{d}\tau}\\ &= \frac{\theta\int e^{-\frac{\mathscr{H}_r}{\theta}}\,\mathrm{d}u_r - \theta\left[u_r\cdot e^{-\frac{\mathscr{H}_r}{\theta}}\right]_{-\infty}^{+\infty}}{\int e^{-\frac{\mathscr{H}_r}{\theta}}\,\mathrm{d}u_r}\,.\end{aligned}$$

But

$$\begin{aligned}\left[u_r\cdot e^{-\frac{\mathscr{H}_r}{\theta}}\right]_{-\infty}^{+\infty} &= \left[\frac{u_r}{\left(e^{-q_ru_r} + q_r\cdot e^{(1-q_r)\cdot u_r}\right)^{\frac{\beta_r}{\theta}}}\right]_{-\infty}^{+\infty}\\ &= 0\end{aligned}$$

Hence

$$\bar{B}_r = \theta\,.$$

(c) The average of the function $A_r^2 = \beta_r^2 \cdot (X_r - q_r)^2$:

$$\begin{aligned}\beta_r^2 \cdot \overline{(X_r - q_r)^2} &= \frac{\int \left(\frac{\partial \mathscr{H}}{\partial u_r}\right)^2 \cdot e^{-\frac{\mathscr{H}}{\theta}} \, \mathrm{d}\tau}{e^{-\frac{\mathscr{H}}{\theta}} \, \mathrm{d}\tau} \\ &= \frac{-\theta \left[\frac{\partial \mathscr{H}_r}{\partial u_r} \cdot e^{-\frac{\mathscr{H}_r}{\theta}}\right]_{-\infty}^{+\infty} + \theta \int e^{-\frac{\mathscr{H}_r}{\theta}} \frac{\partial^2 \mathscr{H}_r}{\partial u_r^2} \, \mathrm{d}u_r}{\int e^{-\frac{\mathscr{H}_r}{\theta}} \, \mathrm{d}u_r} \,.\end{aligned}$$

We also have

$$\frac{\partial \mathscr{H}}{\partial u_r} = \beta_r \cdot \left(\frac{q_r \cdot e^{u_r}}{1 + q_r \cdot e^{u_r}} - q_r\right)$$

so that

$$\frac{\partial^2 \mathscr{H}}{\partial u_r^2} = \beta_r \cdot \left(\frac{q_r \cdot e^{u_r}}{1 + q_r \cdot e^{u_r}}\right) \cdot \left(1 - \frac{q_r \cdot e^{u_r}}{1 + q_r \cdot e^{u_r}}\right) ,$$

i.e.

$$\frac{\partial^2 \mathscr{H}}{\partial u_r^2} = \beta_r q_r \cdot (1 - 2q_r) + (1 - 2q_r) \cdot \frac{\partial \mathscr{H}}{\partial u_r} - \frac{1}{\beta_r} \cdot \left(\frac{\partial \mathscr{H}}{\partial u_r}\right)^2 .$$

Hence

$$\begin{aligned}\left[\frac{\partial \mathscr{H}_r}{\partial u_r} \cdot e^{-\frac{\mathscr{H}_r}{\theta}}\right]_{-\infty}^{+\infty} &= \left[\beta_r \cdot \left(\frac{q_r \cdot e^{u_r}}{1 + q_r \cdot e^{u_r}} - q_r\right) \cdot e^{\frac{-\beta_r}{\theta} \cdot \ln(1 + q_r \cdot e^{u_r}) - q_r u_r}\right]_{-\infty}^{+\infty} \\ &= \left[\beta_r \cdot \left(\frac{1}{\frac{1}{q_r} \cdot e^{u_r} + 1} - q_r\right) \cdot \left(\frac{1}{\frac{1}{e^{q_r u_r}} + q_r \cdot e^{(1 - q_r) \cdot u_r}}\right)\right]_{-\infty}^{+\infty} \\ &= 0 \,.\end{aligned}$$

Thus

$$\begin{aligned}\beta_r^2 \cdot \overline{(X_r - q_r)^2} &= \frac{\theta \int e^{-\frac{\mathscr{H}_r}{\theta}} \frac{\partial^2 \mathscr{H}_r}{\partial u_r^2} \, \mathrm{d}u_r}{\int e^{-\frac{\mathscr{H}_r}{\theta}} \, \mathrm{d}u_r} \\ &= \frac{\theta \int e^{-\frac{\mathscr{H}_r}{\theta}} \cdot \left\{\beta_r q_r \cdot (1 - q_r) + (1 - 2q_r) \frac{\partial \mathscr{H}_r}{\partial u_r} - \frac{1}{\beta_r} \left(\frac{\partial \mathscr{H}}{\partial u_r}\right)^2\right\} \mathrm{d}u_r}{\int e^{-\frac{\mathscr{H}_r}{\theta}} \, \mathrm{d}u_r} \\ &= \frac{\theta \beta_r q_r \cdot (1 - q_r)}{1 + \frac{\theta}{\beta_r}} \qquad (5.8)\end{aligned}$$

and therefore

$$\theta = \beta_r \cdot \frac{\overline{\frac{(X_r - q_r)^2}{q_r \cdot (1-q_r)}}}{1 - \overline{\frac{(X_r - q_r)^2}{q_r \cdot (1-q_r)}}} \tag{5.9}$$

This enables us to give an interpretation of the constant θ, as the "*temperature*" of the neuronal network, for in addition to the above result we also have the equation:

$$\theta = \overline{u_r \frac{\partial \mathscr{H}}{\partial u_r}} = \overline{\beta_r \cdot \ln\left(\frac{\frac{X_r}{q_r}}{1 - X_r}\right) \cdot (X_r - q_r)}\,.$$

The quantity $\bar{B}_r = \overline{u_r \frac{\partial \mathscr{H}}{\partial u_r}}$ is essentially the analogue of the "*kinetic energy*" of the network, and we have the interesting result that *this "kinetic energy" is equipartitioned throughout the network*, since the mean $\bar{B}$ is the same for every neuron equal to the temperature of the network.
Equation 5.9 indicates that *the neuronal temperature is a measure of the mean square deviations of the variables X_r from their stationary states q_r*. Obviously zero temperature corresponds to no deviations from the stationary state.

(d) We now compute the Gibbs phase integral or *Partition function*:

$$Z = \int e^{-\frac{\mathscr{H}}{\theta}}\,\mathrm{d}\tau = \prod_r \int_{-\infty}^{+\infty} e^{-\frac{\mathscr{H}_r}{\theta}}\,\mathrm{d}u_r = \prod_r Z_r$$

Now

$$\begin{aligned}\int_{-\infty}^{+\infty} e^{-\frac{\mathscr{H}_r}{\theta}}\,\mathrm{d}u_r &= \int_{-\infty}^{+\infty} e^{\frac{\beta_r}{\theta}} \cdot \left[\ln\left(1 + q_r \cdot e^{u_r}\right) - q_r u_r\right]\mathrm{d}u_r \\ &= \int_0^1 (1 - X_r)^{\frac{\beta_r}{\theta}\cdot(1-q_r)-1} \cdot X_r^{\frac{\beta_r q_r}{\theta}-1} \cdot q_r^{-\frac{\beta_r q_r}{\theta}}\,\mathrm{d}X_r \\ &= q_r^{-\frac{\beta_r q_r}{\theta}} \cdot B\left(\beta_r q_r/\theta, \beta_r/\theta \cdot (1 - q_r)\right)\,,\end{aligned}$$

where $B(\ ,\ )$ is the well-known Beta-Function of Euler (see [26]). Thus the partition function is given by

$$Z = \prod_r q_r^{-\frac{\beta_r q_r}{\theta}} \cdot B\left(\frac{\beta_r}{\theta}, \frac{\beta_r}{\theta} \cdot (1 - q_r)\right)\,. \tag{5.10}$$

From this we can calculate the probability that one neuron will have its sensitivity

X_r in the range $X_r, X_r + \mathrm{d}X_r$. This is

$$\begin{aligned} P(X_r)\,\mathrm{d}X_r =& \frac{X_r^{\frac{\beta_r q_r}{\theta}-1}\cdot(1-X_r)^{\frac{\beta_r}{\theta}(1-q_r)-1}\cdot q_r^{-\frac{\beta_r q_r}{\theta}}}{q_r^{-\frac{\beta_r q_r}{\theta}}\cdot B\left(\frac{\beta_r q_r}{\theta},\frac{\beta_r}{\theta}\cdot(1-q_r)\right)}\,\mathrm{d}X_r \\ =& \frac{(X_r)^{\frac{\beta_r q_r}{\theta}-1}\cdot(1-X_r)^{\frac{\beta_r}{\theta}(1-q_r)-1}}{B\left(\frac{\beta_r q_r}{\theta},\frac{\beta_r}{\theta}\cdot(1-q_r)\right)}\,\mathrm{d}X_r \end{aligned}$$

We recognize this as the well-known *Beta distribution* of statistical theory (see [79]), usually written in the form:

$$P(X)\,\mathrm{d}X = \frac{1}{B(p,q)}\cdot X^{p-1}(1-X)^{q-1}\,\mathrm{d}X\,, \quad p,q>0\,. \tag{5.11}$$

The range of the distribution is from $X = 0$ to $X = 1$. Its distribution function is the incomplete Beta function (Courant op.cit.). When both p and q exceed unity, the distribution is zero at its extremities and has a unique mode at $(p-1)/(p+q-2)$. Figure C.10 shows some examples of this two parameter distribution.
In our case

$$\begin{aligned} p =& \frac{\beta_r q_r}{\theta} \\ q =& \frac{\beta_r\cdot(1-q_r)}{\theta} \end{aligned}$$

so that $p + q = \beta_r/\theta$.
Thus two conditions, the unimodality of the distribution and the finiteness of θ serve to determine the maximum value of $\overline{X_r^2}$ and hence of the peak to peak amplitude of the variations in X_r about q_r. Using these conditions we obtain the inequality

$$\overline{(X_r-q_r)^2} < \frac{\min(q_r,1-q_r)}{1+\min(q_r,1-q_r)}\cdot q_r\,(1-q_r)\,. \tag{5.12}$$

The right hand side of this equation is zero at its extremities $q_r = 0$ and 1, and takes its maximum value of 1/12 at $q_r = 1/2$. From this we infer that the peak-to-peak amplitude of any oscillator is a maximum whenever the stationary state equals 1/2, and tends to zero as the stationary state approaches 0 or 1. (If either p or q is less than 1, which can occur, the distribution becomes J-shaped or even U-shaped, in which case these conclusions no longer hold, and the peak-to-peak amplitude of the variations can increase near the extremities).
The moments of the Beta distribution are given by

$$\int_0^1 X^k P(X)\,\mathrm{d}X = \frac{\Gamma(p+k)\cdot\Gamma(p+q)}{\Gamma(p)\cdot\Gamma(p+q+1)} \tag{5.13}$$

so that the mean is $p/(p+q)$ and the variance is $pq/(p+q)^2 \cdot (p+q+1)$.
Thus $\overline{X}_r = q_r$ as required. If we compare the mode with the mean of the distribution we see that:

$$q_r \begin{cases} > \frac{(\beta_r q_r/\theta)-1}{(\beta_r/\theta)-2}, & \text{if } q_r < (1-q_r) \\ = \frac{(\beta_r q_r/\theta)-1}{(\beta_r/\theta)-2}, & \text{if } q_r = (1-q_r) \\ < \frac{(\beta_r q_r/\theta)-1}{(\beta_r/\theta)-2}, & \text{if } q_r > (1-q_r)\,. \end{cases} \tag{5.14}$$

Thus when $q_r < 1/2$ the distribution weights values of X_r less than q_r. When $q_r > 1/2$ the distribution weights values of X_r greater than q_r, and when $q_r = 1/2$ the mean and the mode coincide.
Figure C.11 depicts some examples of the oscillations about various values of q_r. It will be seen that the distribution correctly predicts the general nature of the oscillations. In fact we can investigate these more closely by examining the fractions T_-/T and T_+/T of a long interval of duration T spent by an oscillator below and above the mean level q_r.
Consider the function:

$$h_-\left(X_r - q_r\right) = \begin{cases} 1, & \text{if } X_r < q_r \\ 0, & \text{if } X_r > q_r\,. \end{cases} \tag{5.15}$$

Then:

$$\begin{aligned} T_-/T &= \langle h_-\left(X_r - q_r\right)\rangle = E\left\{h_-\left(X_r - q_r\right)\right\} \\ &= \int_{-\infty}^{+\infty} h_-\left(X_r - q_r\right) \cdot e^{-\frac{\mathscr{H}_r}{\theta}}\,\mathrm{d}u_r/Z_r \\ &= \int_0^1 h_-\left(X_r - q_r\right) X_r^{p-1}\left(1-X_r\right)^{q-1}\mathrm{d}X_r/B(p,q) \\ &= \int_0^{q_r} X_r^{p-1}\left(1-X_r\right)^{q-1}\mathrm{d}X_r/B(p,q) \\ &= \frac{B_{qr}(p,q)}{B(p,q)} \\ &= I_{q_r}(p,q)\,, \quad \text{the incomplete Beta density function.} \end{aligned} \tag{5.16}$$

We remark that we could have written down this result immediately, by interchanging time-averages with ensemble-averages, for $I_{q_r}(p,q)$ is of course the probability distribution function $\Pr\left[X_r < q_r\right]$.

It follows that:

$$\begin{aligned} T_+/T =& 1 - T_-/T = \Pr\left[X_r > q_r\right] \\ =& 1 - I_{q_r}(p,q) \\ =& I_{1-q_r}(q,p) \end{aligned} \tag{5.17}$$

We can also compute the mean amplitudes A_- and A_+ of the oscillations below and above q_r. Thus:

$$\begin{aligned} A_- &= \frac{\langle h_-\left(X_r - q_r\right) \cdot \left(X_r - q_r\right)\rangle}{\langle h_-\left(X_r - q_r\right)\rangle} \\ &= \frac{E\left\{h_-\left(X_r - q_r\right) \cdot \left(X_r - q_r\right)\right\}}{E\left\{h_-\left(X_r - q_r\right)\right\}} \\ &= \frac{\int_0^{q_r}\left(X_r - q_r\right) X_r^{p-1}\left(1 - X_r\right)^{q-1}\mathrm{d}X_r}{\int_0^{q_r} X_r^{p-1}\left(1 - X_r\right)^{q-1}\mathrm{d}X_r} \\ &= \frac{B_{q_r}(p+1,q) - q_r B_{q_r}(p,q)}{B_{q_r}(p,q)} \\ &= \frac{B_{q_r}(p+1,q)}{B_{q_r}(p,q)} - q_r \end{aligned}$$

It is easily shown that

$$B_{q_r}(p+1,q) = \frac{p}{p+q} \cdot B_{q_r}(p,q) - \frac{q_r^p \cdot (1-q_r)^q}{(p+q)}$$

and since $p/(p+q) = q_r$ we have

$$A_- = -\frac{q_r^p \cdot (1-q_r)^q}{(p+q)B_{q_r}(p,q)}\,. \tag{5.18}$$

Similarly

$$\begin{aligned} A_+ &= \frac{B_{1-q_r}(q,p+1)}{B_{1-q_r}(q,p)} - q_r \\ &= +\frac{q_r^p \cdot (1-q_r)^q}{(p+q)B_{1-q_r}(q,p)}\,. \end{aligned} \tag{5.19}$$

We can also write down an expression for the mean peak-to-peak amplitude of the

oscillations:

$$A = A_+ - A_- = \frac{q_r^p \cdot (1-q_r)^q}{p+q} \cdot \left\{ \frac{1}{B_{1-q_r}(q,p)} + \frac{1}{B_{q_r}(p,q)} \right\} \tag{5.20}$$
$$= \frac{1}{p+q} \cdot \frac{q_r^p}{I_{q_r}(p,q)} \cdot \frac{(1-q_r)^q}{B_{1-q_r}(q,p)}$$

where we have used the relation $B_\alpha(x,y) = B(x,y) - B_{1-\alpha}(y,x)$.

(e) We now turn to the calculation of an important "*observable*" associated with the network of oscillators: the mean frequency of threshold crossing of the sensitivity function X_r about some fixed threshold value X.
The time-average of this frequency is evidently given by:

$$\begin{aligned}
\overline{w} &= \langle |\dot{X}_r| \delta(X_r - X) \rangle \\
&= E\left\{ |\dot{X}_r| \delta(X_r - X) \right\} \\
&= \frac{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} |\dot{X}_r| \delta(X_r - X) \cdot e^{-\frac{\mathscr{H}_r}{\theta}} \, \mathrm{d}u_1 \, \mathrm{d}u_2 \dots \mathrm{d}u_{2N}}{Z} \\
&= \int_0^1 \int_0^1 \dots \int_0^1 |\dot{X}_r| \delta(X_r - X) \cdot P(X_1, X_2, \dots, X_{2N}) \, \mathrm{d}X_1 \, \mathrm{d}X_2 \dots \mathrm{d}X_{2N} \, .
\end{aligned}$$

But $X_r \in (0,1)$, so that $X_r(1-X_r) > 0$, hence

$$\dot{X}_r = \left| \left(\gamma_r + \frac{1}{\beta_r} \cdot \sum_s \alpha_{sr} X_s \right) X_r \cdot (1 - X_r) \right| = X_r \, (1 - X_r) \left| \gamma_r + \frac{1}{\beta_r} \sum_s \alpha_{sr} X_s \right|$$

and recalling that

$$P\left(X_1, X_2, \dots, X_{2N}\right) = \prod_i P(X_i), \quad P(X_i) = \frac{X_i^{\bar{p}_i - 1} \, (1 - X_i)^{\bar{q}_i - 1}}{B(\bar{p}, \bar{q})} \, ,$$

we have

$$\begin{aligned}
\overline{w} = & \frac{\int_0^1 X_r^{\bar{p}_r} \, (1 - X_r)^{\bar{q}_r} \, \delta\left(X_r - X\right) \mathrm{d}X_r}{B(\bar{p}_r, \bar{q}_r)} \\
& \cdot \int_0^1 \int_0^1 \dots \int_0^1 \left| \gamma_r + \frac{1}{\beta_r} \sum_s \alpha_{sr} X_s \right| \prod_s P(X_s) \, \mathrm{d}X_s
\end{aligned} \tag{5.21}$$

where we have introduced the notation

$$\bar{p}_i = \frac{\beta_i q_i}{\theta}$$
$$\bar{q}_i = \frac{\beta_i \cdot (1 - q_i)}{\theta}$$

in order to avoid confusion.
The first of the integrals is easily evaluated, and gives a contribution

$$\frac{X^{\bar{p}_r} \cdot (1 - X_r)^{\bar{q}_r}}{B(p_r, q_r)} \tag{5.22}$$

towards the product.
The difficult part is calculating the expectation of $\left|\gamma_r + \frac{1}{\beta_r}\sum_s \alpha_{sr} X_s\right|$. We start from the well-known formula[1]:

$$|y| = \frac{1}{\pi} \cdot \int_{-\infty}^{+\infty} \frac{1 - \cos(yt)}{t^2}\,\mathrm{d}t = \mathscr{R}\left(\frac{1}{\pi} \cdot \int_{-\infty}^{+\infty} \frac{1 - e^{i\,ty}}{t^2}\,\mathrm{d}t\right).$$

Hence

$$\begin{aligned}
&\int_0^1 \int_0^1 \cdots \int_0^1 \left|\gamma_r + \frac{1}{\beta_r}\sum_s \alpha_{sr} X_s\right| \prod_s P(X_s)\,\mathrm{d}X_s \\
&= \int_0^1 \int_0^1 \cdots \int_0^1 \mathscr{R}\left(\frac{1}{\pi} \cdot \int_{-\infty}^{+\infty} \frac{1 - e^{i\,t\left(\gamma_r + \frac{1}{\beta_r}\sum_s \alpha_{sr} X_s\right)}}{t^2}\,\mathrm{d}t\right) \prod_s P(X_s)\,\mathrm{d}X_s \\
&= \mathscr{R}\left(\frac{1}{\pi} \cdot \int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{t^2}\left[1 - \int_0^1 \int_0^1 \cdots \int_0^1 e^{i\,t\left(\gamma_r + \frac{1}{\beta_r}\sum_s \alpha_{sr} X_s\right)} \prod_s P(X_s)\,\mathrm{d}X_s\right]\right)
\end{aligned}$$

where we have used the normalizing property of $P(X)$ and the assumed uniform convergence of the integrals. The integral further reduces to:

$$\frac{1}{\pi} \cdot \int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{t^2} \cdot \left[1 - \mathscr{R}\left(e^{i\,t\in_r} \prod_s \int_0^1 e^{i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right)X_s} P(X_s)\,\mathrm{d}X_s\right)\right].$$

But

$$\int_0^1 e^{i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right)X_s} P(X_s)\,\mathrm{d}X_s$$

is the characteristic function of the variable $\left(\frac{\alpha_{sr}}{\beta_r}\right) X_s$, and since $P(X_s)$ is the Delta

[1] EULER's formula: $e^{ity} = \cos ty + i \sin ty$

distribution, we have the well-known result that

$$\int_0^1 e^{i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right)X_s} P(X_s)\,\mathrm{d}X_s = {}_1F_1\left[\bar{p}_s;\bar{p}_s+\bar{q}_s;i\,t\,(\alpha_{sr}/\beta_r)\right]\,, \tag{5.23}$$

where ${}_1F_1\left[\alpha;\beta;X\right]$ is the confluent hypergeometric function (see [143]):

$$\begin{aligned} {}_1F_1\left[a;b;x\right] =& \sum_{n=0}\frac{(a)_n}{(b)_n}\cdot\frac{x^n}{n!}\,, \quad b>0 \\ (a)_n =& a\cdot(a+1)\cdot(a+2)\cdot\ldots\cdot(a+n-1)\,, \quad (a)_0=1\,. \end{aligned} \tag{5.24}$$

Thus our integral reduces to

$$\mathscr{R}\left(\frac{1}{\pi}\cdot\int_{-\infty}^{+\infty}\frac{\mathrm{d}t}{t^2}\cdot\left[1-e^{i\,t\in_r}\prod_s {}_1F_1\left[\bar{p}_s;\bar{p}_s+\bar{q}_s;i\,t\,(\alpha_{sr}/\beta_r)\right]\right]\right)\,.$$

We now recall Equation 4.6a that

$$\gamma_r+\frac{1}{\beta_r}\cdot\sum_s\alpha_{sr}q_s=0$$

so that we have finally to integrate the expression:

$$\mathscr{R}\left(\frac{1}{\pi}\cdot\int_{-\infty}^{+\infty}\frac{\mathrm{d}t}{t^2}\cdot\left[1-\prod_s e^{-i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right)q_s}\;{}_1F_1\left[\bar{p}_s;\bar{p}_s+\bar{q}_s;i\,t\,(\alpha_{sr}/\beta_r)\right]\right]\right)\,.$$

We consider first the case of $s=1$, a neuronal oscillator comprising only two neurons, one excitatory, one inhibitory. We use our knowledge of the fact that since ${}_1F_1\left[\bar{p}_s;\bar{p}_s+\bar{q}_s;i\,t\,(\alpha_{sr}/\beta_r)\right]$ is an analytic function of $t(\alpha_{sr}/\beta_r)$, and since it is the characteristic function of the Beta distribution, its Fourier transform exists, and in fact is obviously the Beta distribution, i.e.:

$$\begin{aligned} &\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right)X_s}\;{}_1F_1\left[\bar{p}_s;\bar{p}_s+\bar{q}_s;i\,t\,(\alpha_{sr}/\beta_r)\right]\mathrm{d}\left(t\,(\alpha_{sr}/\beta_r)\right) \\ &\qquad = \begin{cases} P(X_s), & \text{if } X_s\in(0,1)\,, \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{5.25}$$

We make use of a well-know theorem (see [26]) that

$$\frac{\mathrm{d}}{\mathrm{d}z}\int_a^b f(x,z)\,\mathrm{d}x=\int_a^b\frac{\partial f(x,z)\,\mathrm{d}x}{\partial z} \tag{5.26}$$

in case $f(x,z)$ is a uniformly continuous function of z with continuous derivatives.

We consider the integrand

$$f\left(t,\left(\alpha_{sr}/\beta_r\right)\right)=\frac{1}{t^2}\cdot\left\{1-e^{-i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right)q_s}\ {}_1F_1\left[\beta_s q_s/\theta;\beta_s/\theta;i\,t\left(\alpha_{sr}/\beta_r\right)\right]\right\}$$

recalling that $\bar{p}_s=\beta_s q_s/\theta$, and $\bar{q}_s=\beta_s\cdot\left(1-q_s\right)/\theta$.
Then

$$\begin{aligned}\frac{\partial f(t,\left(\alpha_{sr}/\beta_r\right))}{\partial\left(\alpha_{sr}/\beta_r\right)}&=\frac{i}{t}q_s\cdot e^{-i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right)q_s}\ {}_1F_1\left[\beta_s q_s/\theta;\beta_s/\theta;i\,t\left(\alpha_{sr}/\beta_r\right)\right]\\&\quad-\frac{1}{t^2}\cdot e^{-i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right)q_s}\frac{\partial}{\partial\left(\frac{\alpha_{sr}}{\beta_r}\right)}\ {}_1F_1\left[\beta_s q_s/\theta;\beta_s/\theta;i\,t\left(\alpha_{sr}/\beta_r\right)\right]\ .\end{aligned}$$

Now

$$\frac{\mathrm{d}}{\mathrm{d}x}\left({}_1F_1\left[a;b;x\right]\right)=\frac{a}{b}\cdot{}_1F_1\left[a+1;b+1;x\right]\tag{5.27}$$

so that our derivative equals:

$$\begin{aligned}&\frac{i}{t}q_s\cdot e^{-i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right)q_s}\cdot\left\{{}_1F_1\left[\beta_s q_s/\theta;\beta_s/\theta;i\,t\left(\alpha_{sr}/\beta_r\right)\right]\right.\\&\qquad\left.-{}_1F_1\left[\beta_s q_s/\theta+1;\beta_s/\theta+1;i\,t\left(\alpha_{sr}/\beta_r\right)\right]\right\}\ .\end{aligned}$$

But

$$\begin{aligned}&{}_1F_1\left[\beta_s q_s/\theta;\beta_s/\theta;i\,t\left(\alpha_{sr}/\beta_r\right)\right]-{}_1F_1\left[\beta_s q_s/\theta+1;\beta_s/\theta+1;i\,t\left(\alpha_{sr}/\beta_r\right)\right]\\&=\sum_{n=0}^{\infty}\left\{\frac{\left(\beta_s q_s/\theta\right)_n}{\left(\beta_s/\theta\right)_n}-\frac{\left(\beta_s q_s/\theta+1\right)_n}{\left(\beta_s/\theta+1\right)_n}\right\}\cdot\frac{\left(i\,t\left(\alpha_{sr}/\beta_r\right)\right)^n}{n!}\\&=\sum_{n=1}^{\infty}\frac{\left(\beta_s q_s/\theta+1\right)_{n-1}}{\left(\beta_s/\theta+1\right)_n}\cdot n\left(q_s-1\right)\cdot\frac{\left(i\,t\left(\alpha_{sr}/\beta_r\right)\right)^n}{n!}\\&=\frac{i\,t\left(\alpha_{sr}/\beta_r\right)\left(q_s-1\right)}{\left(\beta_s/\theta+1\right)}\cdot\sum_{n=1}^{\infty}\frac{\left(\beta_s q_s/\theta+1\right)_{n-1}}{\left(\beta_s/\theta+2\right)_{n-1}}\cdot\frac{\left(i\,t\left(\alpha_{sr}/\beta_r\right)\right)^{n-1}}{(n-1)!}\\&=\frac{i\,t\left(\alpha_{sr}/\beta_r\right)\left(q_s-1\right)}{\left(\beta_s/\theta+1\right)}\ {}_1F_1\left[\beta_s q_s/\theta+1;\beta_s/\theta+2;i\,t\left(\alpha_{sr}/\beta_r\right)\right]\ .\end{aligned}$$

Hence our derivative finally reduces to:

$$\frac{q_s\cdot\left(1-q_s\right)}{\left(\beta_s/\theta+1\right)}\cdot\left(\frac{\alpha_{sr}}{\beta_r}\right)e^{-i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right)q_s}\ {}_1F_1\left[\beta_s q_s/\theta+1;\beta_s/\theta+2;i\,t\left(\alpha_{sr}/\beta_r\right)\right]\ .$$

Thus

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\left(\alpha_{sr}/\beta_r\right)} \cdot \left\{ \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{t^2} \left(1 - e^{-i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right) q_s} \; {}_1F_1\left[\beta_s q_s/\theta; \beta_s/\theta; i\,t\left(\alpha_{sr}/\beta_r\right)\right] \right) \right\} \\
&= \frac{1}{\pi} \left(\frac{\alpha_{sr}}{\beta_r} \right) \frac{q_s \cdot (1 - q_s)}{(\beta_s/\theta + 1)} \\
&\quad \cdot \int_{-\infty}^{+\infty} e^{-i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right) q_s} \; {}_1F_1\left[\beta_s q_s/\theta + 1; \beta_s/\theta + 2; i\,t\left(\alpha_{sr}/\beta_r\right)\right] \mathrm{d}t \,.
\end{aligned}$$

An immediate application of Equation 5.25 reduces this to

$$\begin{aligned}
&\frac{1}{\pi} \left(\frac{\alpha_{sr}}{\beta_r} \right) \frac{q_s \cdot (1 - q_s)}{(\beta_s/\theta + 1)} \cdot \frac{2\pi}{(\alpha_{sr}/\beta_r)} \cdot \frac{q_s^{\beta_s q_s/\theta} \cdot (1 - q_s)^{\beta_s (1 - q_s)/\theta}}{B\left(\beta_s q_s/\theta + 1, \beta_s (1 - q_s)/\theta + 1\right)} \\
&= \frac{2}{(\beta_s/\theta + 1)} \cdot \frac{q_s^{\beta_s q_s/\theta + 1} \cdot (1 - q_s)^{\beta_s (1 - q_s)/\theta + 1}}{B\left(\beta_s q_s/\theta + 1, \beta_s \left(1 - q_s\right)/\theta + 1\right)} \,.
\end{aligned}$$

Thus we have

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\left(\alpha_{sr}/\beta_r\right)} \cdot \left\{ \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{t^2} \left(1 - e^{-i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right) q_s} \; {}_1F_1\left[\beta_s q_s/\theta; \beta_s/\theta; i\,t\left(\alpha_{sr}/\beta_r\right)\right] \right) \right\} \\
&= \frac{2}{(\beta_s/\theta + 1)} \cdot \frac{q_s^{\beta_s q_s/\theta + 1} \cdot (1 - q_s)^{\beta_s (1 - q_s)/\theta + 1}}{B\left(\beta_s q_s/\theta + 1, \beta_s \left(1 - q_s\right)/\theta + 1\right)} \,,
\end{aligned} \tag{5.28}$$

so that

$$\begin{aligned}
&\frac{1}{\pi} \cdot \int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{t^2} \left(1 - e^{-i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right) q_s} \; {}_1F_1\left[\beta_s q_s/\theta; \beta_s/\theta; i\,t\left(\alpha_{sr}/\beta_r\right)\right] \right) \\
&= \frac{2}{(\beta_s/\theta + 1)} \cdot \frac{q_s^{\beta_s q_s/\theta + 1} \cdot (1 - q_s)^{\beta_s (1 - q_s)/\theta + 1}}{B\left(\beta_s q_s/\theta + 1, \beta_s \left(1 - q_s\right)/\theta + 1\right)} \cdot \left(\frac{\alpha_{sr}}{\beta_r} \right) + C \,.
\end{aligned}$$

The fact that ${}_1F_1\left[a; b; 0\right] = 1$ allows us to set the constant of integration C equal to zero, and since the result of the integration is a real function, we finally have the result that:

$$\begin{aligned}
&\mathscr{R}\left(\frac{1}{\pi} \cdot \int_{-\infty}^{+\infty} \frac{\mathrm{d}t}{t^2} \left(1 - e^{-i\,t\left(\frac{\alpha_{sr}}{\beta_r}\right) q_s} \; {}_1F_1\left[\beta_s q_s/\theta; \beta_s/\theta; i\,t\left(\alpha_{sr}/\beta_r\right)\right] \right) \right) \\
&= \frac{2}{(\beta_s/\theta + 1)} \cdot \frac{q_s^{\beta_s q_s/\theta + 1} \cdot (1 - q_s)^{\beta_s (1 - q_s)/\theta + 1}}{B\left(\beta_s q_s/\theta + 1, \beta_s \left(1 - q_s\right)/\theta + 1\right)} \cdot \left(\frac{\alpha_{sr}}{\beta_r} \right) \,.
\end{aligned} \tag{5.29}$$

We thus obtain the result that the mean frequency of zeros of the r-th neuron

about some level of sensitivity X is given by:

$$\overline{w\left(X_r - X\right)} = \frac{X^{\beta_r q_r/\theta} \cdot (1-X)^{\beta_r(1-q_r)/\theta}}{B\left(\beta_r q_r/\theta, \beta_r\left(1-q_r\right)/\theta\right)} \cdot \frac{2}{\left(\beta_s/\theta+1\right)} \cdot \frac{q_s^{\beta_s q_s/\theta+1} \cdot (1-q_s)^{\beta_s(1-q_s)/\theta+1}}{B\left(\beta_s q_s/\theta+1, \beta_s\left(1-q_s\right)/\theta+1\right)} \cdot \left(\frac{\alpha_{sr}}{\beta_r}\right) \tag{5.30}$$

But it is easily shown that

$$B\left(\beta_s q_s/\theta+1, \beta_s\left(1-q_s\right)/\theta+1\right) = \frac{\beta_s\theta}{\left(\beta_s/\theta+1\right)} q_s\left(1-q_s\right) \cdot B\left(\beta_s q_s/\theta, \beta_s\left(1-q_s\right)/\theta\right)$$

and recalling Equation 5.11 we are able to reduce Equation 5.30 to the relatively simple equation:

$$\overline{w\left(X_r - X\right)} = X \cdot (1-X) P_r(X) \cdot q_s\left(1-q_s\right) P_s(q_s) \cdot \left(\frac{\alpha_{sr}}{\beta_s\beta_r}\right) \cdot 2\theta\,,$$

Where $P_r(X) = \Pr\left(X_r = X\right)$, $P_s(q_s) = \Pr\left(X_s = q_s\right)$.
It follows by symmetry conditions that

$$\overline{w\left(X_s - Y\right)} = Y \cdot (1-Y) P_s(Y) \cdot q_r\left(1-q_r\right) P_r(q_r) \cdot \left(\frac{-\alpha_{rs}}{\beta_r\beta_s}\right) \cdot 2\theta\,,$$

(recall that $\alpha_{sr} = -\alpha_{rs}$, and that we must have $\overline{w\left(X_r - q_r\right)} = \overline{w\left(X_s - q_s\right)}$.)
Thus we have our final result in case $s = 1$:

$$\overline{w\left(X_r - X\right)} = X \cdot (1-X) P_r(X) \cdot q_s\left(1-q_s\right) P_s(q_s) \cdot \left|\gamma_{sr}\right| \cdot 2\theta\,, \tag{5.31}$$

where $\gamma_{sr} = \alpha_{sr}/\left(\beta_s\beta_r\right)$.
The relative simplicity of the above expression suggests that our method for obtaining it was needlessly complicated. Indeed in this case of $s = 1$, the expectation of $\left|\gamma_r + \frac{1}{\beta_r}\alpha_{sr}X_s\right|$ can be obtained much more simply as follows:
We recall that

$$\gamma_r + \frac{1}{\beta_r}\alpha_{sr}q_s = 0$$

so that

$$\begin{aligned}
&\int_0^1 \left|\gamma_r + \frac{1}{\beta_r}\alpha_{sr}X_s\right| P(X_s)\,\mathrm{d}X_s = \int_0^1 \left|\left(\frac{\alpha_{sr}}{\beta_r}\right)(X_s - q_s)\right| P(X_s)\,\mathrm{d}X_s \\
&= \left(\frac{\alpha_{sr}}{\beta_r}\right)\left(E\left\{h_+\left(X_s - q_s\right)\left(X_s - q_s\right)\right\} - E\left\{1 - h_+\left(X_s - q_s\right)\left(X_s - q_s\right)\right\}\right) \\
&= \left(\frac{\alpha_{sr}}{\beta_r}\right)\cdot\frac{\left(B_{1-q_s}\left(\bar{q}_s, \bar{p}_s + 1\right) - q_s B_{1-q_s}\left(\bar{q}_s, \bar{p}_s\right)\right)}{B\left(\bar{p}_s, \bar{q}_s\right)} \\
&\quad - \frac{\left(B_{q_s}\left(\bar{p}_s + 1, \bar{q}_s\right) - q_s B_{q_s}\left(\bar{p}_s, \bar{q}_s\right)\right)}{B\left(\bar{p}_s, \bar{q}_s\right)}\,.
\end{aligned}$$

It is easily shown however that:

$$\begin{aligned}
&B_{1-q_s}\left(\bar{q}_s, \bar{p}_s + 1\right) - q_s B_{1-q_s}\left(\bar{q}_s, \bar{p}_s\right) - B_{q_s}\left(\bar{p}_s + 1, \bar{q}_s\right) + q_s B_{q_s}\left(\bar{p}_s, \bar{q}_s\right) \\
&= 2\cdot\frac{q_s^{\bar{p}_s}\cdot\left(1 - q_s\right)^{\bar{q}_s}}{\left(\bar{p}_s + \bar{q}_s\right)} \\
&= \frac{2\theta}{\beta_s}\cdot q_s^{\bar{p}_s}\cdot\left(1 - q_s\right)^{\bar{q}_s}
\end{aligned}$$

and so we can write down Equation 5.31 immediately; indeed it is evident that

$$\begin{aligned}
&E\left\{\left|\sum_s\left(\frac{\alpha_{sr}}{\beta_r}\right)\left(X_s - q_s\right)\right|\right\} \\
&= 2E\left\{h_+\left(\sum_s\left(\frac{\alpha_{sr}}{\beta_r}\right)\left(X_s - q_s\right)\right)\cdot\left(\sum_s\left(\frac{\alpha_{sr}}{\beta_r}\right)\left(X_s - q_s\right)\right)\right\} \\
&= -2E\left\{h_-\left(\sum_s\left(\frac{\alpha_{sr}}{\beta_r}\right)\left(X_s - q_s\right)\right)\cdot\left(\sum_s\left(\frac{\alpha_{sr}}{\beta_r}\right)\left(X_s - q_s\right)\right)\right\}
\end{aligned}$$

since

$$E\left\{\left|\sum_s\left(\frac{\alpha_{sr}}{\beta_r}\right)\left(X_s - q_s\right)\right|\right\} = 0\,.$$

Thus we need only consider integrals over the range $(0, \alpha)$ where α is some number in the interval (0,1), and we can in fact obtain "formal" solutions to the general case.
Consider first the case of $s = 2$:

We have

$$
\begin{aligned}
& E\left\{h_{-}\left[\left(\frac{\alpha_{1r}}{\beta_r}\right)(X_1-q_1)+\left(\frac{\alpha_{2r}}{\beta_r}\right)(X_2-q_2)\right]\right. \\
& \left.\cdot\left[\left(\frac{\alpha_{1r}}{\beta_r}\right)(X_1-q_1)+\left(\frac{\alpha_{2r}}{\beta_r}\right)(X_2-q_2)\right]\right\} \\
= & \left(\frac{\alpha_{1r}}{\beta_r}\right)\cdot\int_0^{\phi_{11}}\int_0^{\phi_{21}}(X_1-q_1)\,P(X_1)P(X_2)\,\mathrm{d}X_1\,\mathrm{d}X_2 \\
& +\left(\frac{\alpha_{2r}}{\beta_r}\right)\cdot\int_0^{\phi_{12}}\int_0^{\phi_{22}}(X_2-q_2)\,P(X_1)P(X_2)\,\mathrm{d}X_1\,\mathrm{d}X_2
\end{aligned}
$$

where

$$
\begin{aligned}
\phi_{11} &= \frac{\alpha_{1r}q_1+\alpha_{2r}q_2}{\alpha_{1r}} \\
\phi_{21} &= \frac{\alpha_{1r}}{\alpha_{2r}}\cdot(\phi_{11}-X_1) \\
\phi_{12} &= \frac{\alpha_{2r}}{\alpha_{1r}}\cdot(\phi_{22}-X_2) \\
\phi_{22} &= \frac{\alpha_{1r}q_1+\alpha_{2r}q_2}{\alpha_{2r}}\,.
\end{aligned}
$$

That is, we are integrating over the triangular region bounded by the loci $X_1=0$, $X_2=0$, and $(\alpha_{1r}/\beta_r)(X_1-q_1)+(\alpha_{2r}/\beta_r)(X_2-q_2)=0$.
But

$$
\begin{aligned}
& B(\bar{p}_1,\bar{q}_1)\int_0^{\phi_{11}}(X_1-q_1)\,P(X_1)\,\mathrm{d}X_1 = B_{\phi_{11}}(\bar{p}_1+1,\bar{q}_1)-q_1B_{\phi_{11}}(\bar{p}_1,\bar{q}_1) \\
= & -\frac{(\phi_{11})^{\bar{p}_1}\cdot(1-\phi_{11})^{\bar{q}_1}}{(\bar{p}_1+\bar{q}_1)} \\
= & \frac{\theta}{\beta_1}\cdot(\phi_{11})^{\bar{p}_1}\cdot(1-\phi_{11})^{\bar{q}_1}\ .
\end{aligned}
$$

Similarly

$$
B(\bar{p}_2,\bar{q}_2)\int_0^{\phi_{22}}(X_2-q_2)\,P(X_2)\,\mathrm{d}X_2 = \frac{\theta}{\beta_2}\cdot(\phi_{22})^{\bar{p}_2}\cdot(1-\phi_{22})^{\bar{q}_2}
$$

and since ϕ_{11} and ϕ_{22} are independent of X_1 and X_2 we have

$$\begin{aligned}
&\int_0^{\phi_{11}} \int_0^{\phi_{21}} (X_1 - q_1)\, P(X_1)P(X_2)\, \mathrm{d}X_1\, \mathrm{d}X_2 \\
&= \int_0^{\phi_{21}} -\frac{\theta}{\beta_1} \cdot \frac{(\phi_{11})^{\bar{p}_1} \cdot (1-\phi_{11})^{\bar{q}_1}}{B\left(\bar{p}_1, \bar{q}_1\right)} P(X_2)\, \mathrm{d}X_2 \\
&= -\frac{\theta}{\beta_1} \cdot \frac{(\phi_{11})^{\bar{p}_1} \cdot (1-\phi_{11})^{\bar{q}_1}}{B\left(\bar{p}_1, \bar{q}_1\right)} \int_0^{\phi_{21}} P(X_2)\, \mathrm{d}X_2 \\
&= -\frac{\theta}{\beta_1} \cdot \frac{(\phi_{11})^{\bar{p}_1} \cdot (1-\phi_{11})^{\bar{q}_1}\, B_{\phi_{21}}\left(\bar{p}_2, \bar{q}_2\right)}{B\left(\bar{p}_1, \bar{q}_1\right) \cdot B\left(\bar{p}_2, \bar{q}_2\right)}\,.
\end{aligned}$$

Similarly

$$\begin{aligned}
&\int_0^{\phi_{12}} \int_0^{\phi_{22}} (X_2 - q_2)\, P(X_1)P(X_2)\, \mathrm{d}X_1\, \mathrm{d}X_2 \\
&= -\frac{\theta}{\beta_2} \cdot \frac{(\phi_{22})^{\bar{p}_2} \cdot (1-\phi_{22})^{\bar{q}_2}\, B_{\phi_{12}}\left(\bar{p}_1, \bar{q}_1\right)}{B\left(\bar{p}_2, \bar{q}_2\right) \cdot B\left(\bar{p}_1, \bar{q}_1\right)}\,.
\end{aligned}$$

Using these results, simplifying and collecting terms, we obtain the result:

$$\begin{aligned}
\overline{w\left(X_r - X\right)} = X(1-X)P_r(X) \cdot 2\theta\ \{&|\gamma_{1r}| \cdot \phi_{11}\left(1-\phi_{11}\right) P_1\left(\phi_{11}\right) \cdot I_{2,\phi_{21}} \\
&+ |\gamma_{2r}| \cdot \phi_{22}\left(1-\phi_{22}\right) P_2\left(\phi_{22}\right) \cdot I_{1,\phi_{12}}\}\,,
\end{aligned} \tag{5.32}$$

where $I_{s,\phi_{sr}}$ is the incomplete Beta distribution taken w.r.t. to the s-th variable. Similarly we can write down the answer in case $s = 3$:

$$\begin{aligned}
\overline{w\left(X_r - X\right)} = X(1-X)P_r(X) \cdot 2\theta\ \Bigg\{&|\gamma_{1r}| \cdot \phi_{11}\left(1-\phi_{11}\right) P_1\left(\phi_{11}\right) \\
&\cdot \int_0^{\phi_{31}} I_{2,\phi_{21}} P(X_3)\, \mathrm{d}X_3 \\
&+ |\gamma_{2r}| \cdot \phi_{22}\left(1-\phi_{22}\right) P_2\left(\phi_{22}\right) \cdot \int_0^{\phi_{32}} I_{1,\phi_{12}} P(X_3)\, \mathrm{d}X_3 \\
&+ |\gamma_{3r}| \cdot \phi_{33}\left(1-\phi_{33}\right) P_3\left(\phi_{33}\right) \cdot \int_0^{\phi_{13}} I_{2,\phi_{23}} P(X_1)\, \mathrm{d}X_1 \Bigg\}
\end{aligned}$$

Recognizing that

$$\begin{aligned}
\int_0^{\phi_{ij}} I_{k,\phi_{kj}} P(X_i)\, \mathrm{d}X_i &= \int_0^{\phi_{ij}} \int_0^{\phi_{kj}} P(X_i)P(X_k)\, \mathrm{d}X_i\, \mathrm{d}X_k \\
&= \Pr\left\{X_i < \phi_{ij}, X_k < \phi_k\right\}
\end{aligned}$$

and extending the results for general s, we have finally:

$$\boxed{\overline{w\left(X_r - X\right)} = X(1-X)P_r(X)\cdot 2\theta\cdot\left[\sum_s |\gamma_{sr}|\cdot\phi_{ss}\left(1-\phi_{ss}\right)P_s\left(\phi_{ss}\right)\right.}$$
$$\left.\cdot\Pr\left\{X_1 < \phi_{1s}, X_2 < \phi_{2s}, \ldots, X_t, \phi_{ts},\ t\neq s\right\}\right] \tag{5.33}$$

where

$$\phi_{ss} = \sum_t \frac{\alpha_{tr}q_r}{\alpha_{sr}}$$
$$\phi_{ts} = \left(\frac{\alpha_{(t-1)\cdot r}}{\alpha_{tr}}\right)\cdot\left(\phi_{(t-1)\cdot r} - X_{t-1}\right)\ .$$

Equation 5.33 then, represents the most explicit formula that we have so far obtained for the mean frequency of zero (or X) crossings of our neuronal oscillator. By virtue of the close relationship existing between incomplete Beta functions and cumulative binomial distributions we can in fact re-express the above expression in terms of cumulative multinomial distributions.
We conclude this section on threshold crossings by noting that the frequency of crossings of the value X, relative to the frequency of crossings of the value q_r is easily expressed as:

$$w_{\text{rel}}(X) = \frac{\overline{w\left(X_r - X\right)}}{\overline{w\left(X_r - q_r\right)}} = \frac{X(1-X)P(X)}{q_r(1-q)P(q_r)} = \frac{X^{\frac{\beta_r q_r}{\theta}}(1-X)^{\frac{\beta_r(1-q_r)}{\theta}}}{q_r^{\frac{\beta_r q_r}{\theta}}(1-q_r)^{\frac{\beta_r(1-q_r)}{\theta}}}\ . \tag{5.34}$$

(f) Observable coupling parameters
Since $\overline{(X_r - q_r)} = 0$, the sum

$$\sum_s \gamma_{sr}\beta_s\beta_t\overline{(X_s - q_s)(X_t - q_t)} = \gamma_{tr}\beta_t^2\overline{(X_t - q_t)^2}$$

so that we may write the coupling coefficient γ_{tr} as:

$$\gamma_{tr} = \frac{\sum_s \gamma_{sr}\beta_s\beta_t\overline{(X_s - q_s)(X_t - q_t)}}{\beta_t^2\overline{(X_t - q_t)^2}} = \frac{\overline{\frac{1}{X_r(1-X_r)}\cdot X_r\left(X_t - q_t\right)}}{\beta_t\cdot\overline{(X_t - q_t)^2}}\ . \tag{5.35}$$

Thus *we have obtained a means of "observing" the microscopic coupling parameters of our theory.*

An alternative expression for γ_{tr} can also be obtained that involves the temperature θ.

$$\begin{aligned}\gamma_{tr} &= \frac{1}{\theta} \cdot \overline{\sum_s \gamma_{sr}\beta_s(X_s - q_s) \cdot \ln\left(\frac{\frac{X_t}{q_t}}{1 - X_t}\right)} \\ &= \frac{1}{\theta} \cdot \overline{\frac{1}{X_r\,(1 - X_r)}\dot{X}_r \cdot \ln\left(\frac{\frac{X_t}{q_t}}{1 - X_t}\right)}\,. \end{aligned} \tag{5.36}$$

So we can measure the microscopic coupling parameters in terms of the temperature.

(g) Temperature as an observable:
We can also have a means of estimating the neuronal temperature in terms of the relative mean frequency of threshold crossings, for Equation 5.34 can be rewritten as:

$$\theta = \frac{\beta_r q_r \cdot \ln\left(\frac{X}{q_r}\right) + \beta_r(1 - q_r) \cdot \ln\left(\frac{1-X}{1-q_r}\right)}{\ln(w_{\mathrm{rel}}(X))} \tag{5.37}$$

Thus we have θ or θ/β_r in terms of q_r, X, and $w_{\mathrm{rel}}(X)$. Moreover this determination does not depend directly on s, the number of "interactions" per neuron, so that we have a very useful population measure at our disposal, i.e., *neuronal temperature is a real observable of the theory.*
We conclude this chapter by summarizing the results we have obtained by using the Gibbs ensemble. We have shown that the ensemble and time averages of X_r, the sensitivity of the r-th neuron, are equal to the stationary states q_r of the dynamical formulation. We have shown that there is an analogue of kinetic energy for the network, which is equipartitioned throughout all the neurons in the network, and therefore we can introduce the useful notion of temperature for the network as a whole. This "neuronal temperature" was also related to the first two moments of the probability distribution $P(X_r)\,\mathrm{d}X_r$ of the sensitivities, which was shown to be the well-known Beta distribution of mathematical statistics. This distribution was then used to calculate the various fractions of time spent above or below the stationary states q_r by the sensitivities X_r, the peak-to-peak amplitudes of the variations in sensitivity, and the mean frequency of threshold crossings of given values of the sensitivity. Finally it was noted how these measurable quantities allowed inferences to be made about the coupling coefficients α_{rs} of the interaction between neurons, and also how the temperature itself could be estimated from microscopic measurements.

6. Some Applications to Experimental Data

We now wish to show the formalism we have invented can be used to describe, predict and explain some neural phenomena. It is obvious that there is a tremendous mass of experimental data concerning nervous activity that is within the range of this formalism. We do not undertake this here, but only wish to sketch some of the applications, for it has been the main intention of this thesis only to create the necessary formalism. Before we go into the data however, some additional remarks are perhaps necessary concerning the validity of the Beta distribution formalism. As we have already noted in chapter 3, there are only certain ranges wherein constants of the motion can exist in the dynamical formulation. What this means for the statistical mechanics is that *only a certain "spectrum" of temperatures is possible*. We reach this conclusion from the following considerations. The neuronal networks which we are studying are supposedly immersed in a "heat-bath" of many other networks, with which they are supposed to be weakly interacting. This is where the Hamiltonian and Gibbsian formalisms come in. At the same time the networks are receiving stimuli, either from controlling networks elsewhere in the nervous system, or else directly from the "external" world through sense organs and transducers. *The effects of these stimuli are to produce changes in the stationary states q_r of the networks*, for the stimuli are represented by the variables K_r of Equation 3.8. As we have seen in chapter 5, there is a definite relationship between θ, the temperature or mean kinetic energy associated with each variable, and the stationary states q_r. In fact it follows immediately from Equation 5.9 and Equation 5.12, when $P(X_r)\,\mathrm{d}X_r$ is unimodal and well-behaved, that:

$$\theta < \beta_r \ (\min\ q_r, 1 - q_r) \tag{6.1}$$

(If $P(X_r)\,\mathrm{d}X_r$ is J-shaped or U-shaped one has to be a little more circumspect). Evidently θ is greater when $q_r = 1/2$, and falls off as q_r approaches 1 or 0. Now the considerations of chapter 3 imply that only when q_r falls within the intervals where the linear approximation to $\ln\left(\frac{x}{1-x}\right)$ holds, can there be conservation of "energy" and a neuronal temperature. Together with Equation 6.1 this implies *that there should be gaps in the spectrum of temperatures associated with the firing patterns of neuronal networks.*

(1) The validity of the Logistic curve:
A key assumption in our theory is the form of the relationship between $v(t)$, the mean rate at which a neuron emits impulses, and the mean current $I(t)$ driving the neuron. We assumed the relationship to be given by equation Equation 2.6. We may combine equations Equation 2.6 and Equation 2.7 to give:

$$\ln\left(\frac{X_r}{1-X_r}\right) = -\gamma I + \beta \tag{6.2}$$

where $\gamma = \beta/I_{1/2}$. Thus we see that the function $X/(1-X)$, which roughly speaking estimates the "strength" of the input (see the discussion following Equation 2.12) driving the neuron, is proportional to the exponent of the depolarizing current. Is there an experimental test of this relationship? The answer is yes, there are in fact two independent series of experiments which support the assumption. The first series was carried out by O. D. Creutzfeldt et.al. [32] using intracellular microelectrodes to sample both the spikes and intracellular potentials associated with neurons in the Cat Pyramidal Tract; the second by A. D. J. Robertson [129] using extracellular microelectrodes. Figure C.9(a) is a reproduction of Creutzfeldt's data for one Pyramidal cell, and Figure C.12b shows $\ln\left(\frac{X}{1-X}\right)$ plotted as a function of the depolarizing current. It will be seen that the fit to Equation 6.2 is very good. For the extracellular experiments we cannot proceed so directly. Figure C.42 shows the relation of the unitary responses of a neuron in the Cat visual cortex, to the slow potentials recorded on the same extracellular microelectrode in response to light flashes given to the animal once a second (Robertson op.cit.), (see also Figure C.41). The unit response is a Post-Stimulus-Time-Histogram, i.e., it gives the number of spikes elicited by the stimulus, in the intervals k_i following the stimulus, each of duration δr milliseconds. If there are n_i spike counts in the i-th bin, then the "counting rate" (see chapter 2) is given by $n_i/N\delta\tau$, where N is the number of stimulus presentations, and so we take the sensitivity function X to be:

$$X_i = \frac{1 - rn_i}{N\delta\tau} \tag{6.3}$$

for the i-th bin. It will be seen that we have to choose a refractory period. For cortical cells, 1 ms seems to be about the minimum interval observed between spikes, and so we take $r = 1$ ms.
We also have to obtain the depolarizing current from the recorded slow potential. It is well-known that a neuronal membrane is essentially capacitative (see Appendix B). We therefore assume the simplest possible relation, that the *membrane current is proportional to the slope of the observed slow Potential.* We have therefore carried out numerical differentiation on the evoked potential, and then plotted $\ln\left(\frac{X_i}{1-X_i}\right)$ as a function of the differentiated potential, for each "bin" following the stimulus. The results are shown in Figure C.13, for a single cortical cell. Again

it will be seen that the linear fit is good. Several other cells have also given good fits, so that we feel that the status of the Logistic curve in this context is good. This is not to say that other sigmoidal functions would not give better fits, but only that the Logistic function adequately represents the data. (Recently, A. D. J. Robertson, A. Bruner and the author have carried out an extensive series of studies of Cortical cells. Preliminary results indicate that the above considerations are essentially correct. The results will shortly be submitted for publication).

(2) The relation of the frequencies of the nonlinear oscillators to their stationary states:

We turn next to a somewhat different aspect of our theory. It is clear from the various considerations in chapter 3 and chapter 5 that *the frequencies of the nonlinear oscillators* we have constructed are generally speaking *monotonically related to the function* $q_r \cdot (1 - q_r)$. (See for example Equation 5.31). But $q_r \cdot (1 - q_r)$ is a convex function with a maximum value of $1/4$ when $q_r = 1/2$, tending to zero as q_r tends to zero or one. Thus the frequencies are maximal about the stationary states $q_r = 1/2$, and drop to zero for values of q_r tending to 0 or to 1. Now by Equation 2.8' we have $X_r = 1 - \frac{r}{\tau}$, so that $q_r = 1 - \frac{r}{\tau_{av}}$, where τ_{av} is either the ensemble average or the time average of the interspike interval distribution associated with a single component of an oscillating network. We therefore have the relation:

$$\text{Frequency of oscillations } \propto \left(\frac{r}{\tau_{av}}\right) \cdot \left(1 - \frac{r}{\tau_{av}}\right) \tag{6.4}$$

Thus *the frequency of oscillations should increase with* $\frac{1}{\tau_{av}}$ *until the value* $\tau_{av} = 2r$ *is reached, after which the frequency should decrease with* $\frac{1}{\tau_{av}}$. These conclusions obviously follow from the fact that the slope of the Logistic function increases with I until the value $v = 1/2r$ is reached, after which the slope decreases to zero with I. A linear function on the other hand would result in a constant frequency of oscillations, and the exponential function $\frac{1}{r} \cdot \left(1 - e^{-\beta \cdot \frac{I}{I_0} - 1}\right)$ (see Figure C.9) would result only in a decrease of frequency with $\frac{1}{\tau_{av}}$.

It is difficult to find experimental evidence relating to oscillating neuronal networks especially when $\tau_{av} < 2r$, but there is one series of experiments which gives a direct test of our considerations for values of $\tau_{av} > 2r$. This is a series of experiments by G. F. Poggio and L. J. Viernstein, [117], on the firing patterns of thalamic somatic sensory neurons. These neurons are "tonic", i.e., they exhibit maintained firing for long periods of time, and so fall within the range of our theory. They respond to signals from joint-angle receptors in such a way that τ_{av} changes, *and so does the mean frequency of the variations about* τ_{av}. Figure C.14 shows expectation density functions (equivalent to our function $n_i/N\delta\tau$, the "counting-rate" but conditional on previous firings) for a single thalamic somatic sensory neuron, for four different values of joint-angle. It will be seen that the oscillations look like those generated by our equations – flat-troughed with peaks, changing with increasing frequency

into rather more symmetric oscillations, more characteristic of linear oscillations. And of course *the frequency of the oscillations increases with the mean firing rate of the cell*. Thus qualitatively we can say that the general nature of these oscillations in the firing patterns of such neurons does not contradict the kind of theoretical structure we have constructed. Quantitatively however the fit of the frequency of oscillations as a linear function of $q_r \cdot (1 - q_r)$ is not all that good. Figure C.15 shows the results of our calculations, assuming a refractory period of $r = 1\,\text{ms}$. There is a 'kink' in the curve. Of course the dependence on $q_r \cdot (1 - q_r)$ is only one part of Equation 5.31, and we should have to take account of all the variables in the relationship, to obtain accurate quantitative predictions at this level.
These considerations concerning the sigmoidal curve as a basis for the description of input-output behaviour are of course only a small fraction of what can be expected from the theory we have constructed. All the formulas in chapter 5 are expressed in terms of the sensitivity function X_r. But using Equation 2.8' we could translate them all into formulae concerned with interspike intervals. So we could obviously obtain predictions concerning for example, the nature of the interval distributions associated with cells in various conditions. It is easily shown that the interval distribution that results from Equation 5.11 is given by:

$$P(\tau_r)\,\mathrm{d}\tau_r = \frac{\left(1 - \frac{r}{\tau_r}\right)^{\frac{\beta_r q_r}{\theta} - 1} \cdot \left(\frac{r}{\tau_r}\right)^{\frac{\beta(1-q_r)}{\theta} + 2}}{r \cdot B\left(\beta_r q_r/\theta, \beta_r(1-q_r)/\theta\right)}\,\mathrm{d}\tau_r \tag{6.5}$$

and that for example the variance of the interval distribution σ_r^2 is related to the mean interval through the formula:

$$1 + \frac{\beta_r}{\theta} = \frac{\overline{X}_r - \overline{X}_r^2}{\overline{X^2}_r - \overline{X}_r^2} \tag{6.6}$$

subject to the transformation from sensitivity into interval.
It is obvious that there are enough free parameters in the distribution to enable fitting to many of the recent experimental results (see for example [47], [147], [46], [71]). Moreover through Equation 6.1, *mutates mutandis* it is possible to obtain amplitude distributions for the membrane currents and voltages that drive neurons, and to test these against experimental data. In this connexion it is worth noting the following. It is well-known that *if X_r is Beta distributed, then the function $(1-X)/X$ is Pareto distributed*. (see [41], page 47). But $(1-X)/X$ is the function which estimates the "strength" of the inputs which drives the neuron, and as we have seen, is proportional to the exponent of the depolarizing current $I(t)$. Thus we have the immediate result that, effectively, e^I is a Pareto distributed random variable, whenever $X = 1 - \frac{r}{\tau}$ is Beta distributed. Since the Pareto distribution is *stable*, in the terminology of probability theory (see [41], page 165), this means that linear combinations of independent Pareto variables are again (to within a linear

transformation) Pareto distributed. We can therefore use this property to generate predictions concerning the relation of the gross potentials, the ECoG and EEG, to the firing patterns of single neurons. We do not pursue any of these considerations here however, but instead turn to the final piece of experimental data which we promised to consider, the temperature spectrum.

(3) Gaps in the temperature spectrum:

As we have already discussed we do not expect to find that all possible temperatures are well represented in the spectrum of possible temperatures associated with the activity of neurons that form part of our oscillating networks. How are we to test this hypothesis? The answer is that we need only to look at existing data on Post Stimulus Time Histograms (Robertson op.cit.), but we have to take account of the second moments of the distribution of responses, as well as the first moments (see Equation 5.9). We have therefore carried out the following analysis on the responses associated with several cortical neurons to repetitive light flashes. In addition to the sensitivity function X_i associated with each bin (see Equation 6.3), we have calculated the mean sensitivity q_i (this is really what we did before in looking at the validity of the Logistic curve), and also the variance $\overline{(X_i - q_i)^2}$. From these calculations we obtain an estimate of the temperature associated with each bin. (This is well-defined in terms of the ensemble created by the large number of stimulus presentations for each bin). Call it θ_i. We do this for 500 ms following the stimulus, and we also look backwards at the 500 ms preceding the stimulus. The stimuli come along once a second. This gives us a way of comparing immediate with remote effects of the stimulus. For both the forward and backward cases we then calculate and plot histograms of the temperature distribution. The results are shown in Figure C.16. The histograms are actually of the function $Z = 1 + \frac{\beta_r}{\theta}$ defined in Equation 6.6. It will be seen that in all cases there is a pronounced peak at the low temperature end of the histogram (high Z); in the backward cases there are always gaps between this peak and a broader range of higher temperatures, whereas in the forward cases the gaps tend to be somewhat filled, although the pronounced low temperature peak is still present, although smaller, the general effect of the stimulus being a lowering of the temperature distribution. Although the effects are not clear cut, we interpret them as follows. The effect of a stimulus on such cortical neurons is to some extent to diminish their variability (hence the drop in temperature), and to perturb their firing patterns in a marked fashion. However after several hundred milliseconds, the neurons probably reflect the network properties again, rather than those of the stimulus, and it is then that the gaps in the temperature spectrum show up. Although this argument is very rough it does suggest that there are perhaps preferred states of activity in neuronal networks within the central nervous system, that the interaction between such "states" and stimuli is somewhat complicated, but that the neuronal temperature function which we have invented might eventually serve to characterize both the

states and the interactions with stimuli. There is in fact one other piece of evidence which bears on this topic. In a recent paper, G. K. Smith and D. R. Smith, [145], concluded that for many cells in widely scattered regions of the cerebral cortex, the statistics of their interval distributions were best fitted by a compound probability model. Figure C.17 shows one of their histograms. They showed that a good fit to the data was obtained from a Compound Poisson Process (see [29]) in which one "slow" Shot Noise Process is switched at random intervals into another "fast" Shot Noise Process. They visualized the neuronal signal as consisting in part of *bursts* of randomly occurring spikes, the bursts being of unknown duration and frequency. They suggested that one random pulse shower *gated on and off* over shot noise, or alternatively, that it gated on and off over a silent background. They suggested that the random pulse showers might arise in subcortical cells, and noted that the cells they recorded from had to have two "thresholds", one controlling the onset and duration of bursts, the other controlling the actual onset of spikes within bursts. They also suggested that a basis for this condition might lie in the distinctly different configurations of somatic and dendritic synapses that control and drive cortical neurons. It is not difficult to see the relevance of their data and ideas to our nonlinear oscillators. For *nonlinear oscillations in sensitivity correspond to bursts of spikes*. Our considerations imply that there are perhaps two or even *three* preferred states of activity of our oscillators: a low rate-low variability state, a higher rate-high variability state, and a highest-rate low variability state. The first two of these states could correspond to the slow and fast processes seen by Smith and Smith. We know from personal communication ([129]) that there is probably a very fast process corresponding to our third state. If this be so, then the properties of the nonlinear oscillators we have constructed are evidently sufficient to account for this "quantization" into separate and distinct states of the activity of cortical neurons. Although we have not in fact stressed the point, it is well-known ([36]) that interneurons synapse with cortical pyramids, axo-somatically, whereas the remote thalamic inputs from nonspecific nuclei synapse axo-dendritically with the cortical pyramids. As we have seen it is the combination of the interaction of cortical pyramids, cortical interneurons, and remote thalamic cells which is supposed to produce the oscillations in sensitivity of the various cells, and the preferred states of such activity. Because of the relationship between sensitivity and interval (Equation 2.8') it is clear that it would be very difficult to distinguish between combinations of transformed Beta distributions, and the Compound Poisson Process used by Smith and Smith. Thus we feel confident in asserting that the theory we have developed is capable of explaining the findings of Smith and Smith, and others, in terms of fairly simple ideas involving nonlinearities, delays, and feedbacks. On this basis the gaps in the temperature spectrum may well be real, and correspond to the quantization implicit in the above discussion.

7. Summary. Relation to other Work. Concluding Remarks

In summary what we have done is as follows. We have introduced a certain nonlinear delay operator to represent the transfer characteristics of neurons. This operator is not the Heaviside step-function that represents the all-or-none responses of neurons, but is instead a continuous function that effectively represents the mean neuronal response to excitation. So we deal not with the mathematics of switching networks, Boolean algebra and the like, but with differential equations.
We have examined the dynamics of model circuits, certain of which were shown to exhibit undamped responses to stimuli, but for only a few levels of activity. The techniques of Hamiltonian and of Gibbsian statistical mechanics were used to connect these models with experimental data. In doing so we have found the beginnings of an explanation for the existence of "preferred states" in the firing patterns of cortical neurons.
It is clear however, that we have only just scratched the surface of many of the theoretical questions that have to be formulated and answered for this approach to prove really useful. We have certainly oversimplified by neglecting such important aspects of neuronal behaviour as adaptation and accommodation to stimuli, and the changes in latency that are associated with responses to "strong" and "weak" stimuli. One very important simplification we have made is that the input membrane is linear, so that the membrane current built up in the neuron is a linear combination of the responses of all the various sub-synaptic regions of dendritic membrane. This is only true for synapses which are on separate dendritic trees (see [120]). Therefore our basic interaction equation (Equation 2.1) is not true for synapses on the same dendritic tree. What this means is that the Hamiltonian which we have obtained is only the "zeroth" approximation, so to speak, to the very nonlinear Hamiltonian that must serve to describe neuronal interactions. Indeed because of this assumption of linearity, the Hamiltonian (Equation 4.5) was a "sum-function", representing an assembly of (sic) "weakly interacting" neurons, a kind of "*neuron gas*". We are saved from purely gas like behaviour by the delays implicit in the formulation, but nevertheless it is quite clear that to obtain really interesting co-operative effects we shall have to look at nonlinear synaptic interactions in considerable detail. The relation between our work and that of B. G. Cragg and H. N. V. Temperley, [30], remains to be determined, for we have not yet begun to study the *Free Energy function* ψ, that furnishes information about co-operative phenomena.

We have also ignored the very important anatomical or "spatial" aspects of neuronal networks. We have chosen to study only the temporal aspects of certain almost "one-dimensional" corticothalamic circuits, on the grounds that this vertical aspect of neuronal organization is the only aspect that seems to be homologous throughout the central nervous system including the Cerebellum. However it is known that the neuronal networks of the central nervous system have at least a two-dimensional topology. Beurle ([12]) in fact included spatial factors in his treatment of the responses of a randomly connected network of model neurons, in a way that is very closely related to our work. We therefore give a short summary of some aspects of Beurle's treatment.
The neurons are assumed to be randomly distributed with a volume density p. A fraction $R(\underline{r}, t)p$ of them are *sensitive* and the remainder are refractory. The neurons are assumed to have a refractory period r, a threshold q, a delay τ, and a rectangular EPSP $\chi(t)$ of duration s. There is a connexion function $\xi(x)$ defined as the mean number of synapses on a given neuron, from all neurons in an infinite plane of unit thickness, x units away. The fraction of cells becoming active per unit time in any region $F(\underline{r}, t)$ is then introduced, and one-dimensional spatial behaviour is then studied. Beurle showed that the rate at which sensitive neurons become active in the network is proportional to $F(x, t)$, i.e.,that:

$$\frac{\mathrm{d}F(t)}{\mathrm{d}t} =\in R \cdot F(t) \tag{7.1}$$

where the coefficient $\in$ is related to χ and to ξ. For steady state activity in the network, the rate at which sensitive neurons become active should equal the rate at which used neurons recover, i.e.,

$$p \cdot (1 - R) = r \cdot F \cdot p\,. \tag{7.2}$$

If we combine these equations we obtain the equation:

$$\frac{\mathrm{d}R}{\mathrm{d}t} = - \in R \cdot (1 - R) \tag{7.3}$$

or equivalently:

$$\frac{\mathrm{d}}{\mathrm{d}t} \ln \left(\frac{R}{1 - R} \right) = - \in \tag{7.4}$$

We recognize this as our basic differential Equation 3.24 with R analogous to X_r and the coefficient $- \in$ analogous to $\gamma_r + \frac{1}{\beta_r} \cdot \sum_s \alpha_{sr} X_s$. Indeed the analogy is very close, for we may rewrite Equation 7.2 as:

$$R = 1 - r \cdot F \tag{7.5}$$

and compare it with Equation 2.8. Evidently F is analogous to the variable $v(t)$ in our formulation, the mean rate of emission of impulses by a single neuron. Thus if we change our scale so to speak, and look not at the fraction of time that a single unit can be fired, but at the fraction of neurons in a population that can be fired at a single epoch, we apparently end-up with the same differential equation[1]. (See also our discussion on Pareto variables.)

This means that we can extend most of our formalism to deal with at least one spatial dimension, and so start to look at the interactions between neuronal networks on a much more realistic basis. Eventually we hope to have a field theory that will do justice to some aspects of the problem. Of course the considerations concerning damping will again be relevant. Even more so for the coefficient involves the connexion and excitation functions χ and ψ as *convolutions* implying that there is already spatial and temporal *filtering* in the networks. Perhaps our work is really a very specialized case of N. Wiener's general approach to nonlinear filters and to neurological clocks (see [167]). We do not yet know. It is however of considerable interest that I. Prigogine's recent work on non-equilibrium statistical mechanics and thermodynamics (see [119]) indicates that delays and memory are crucial for the maintenance of ordered structures in physical and biological systems, and that the existence of both conservative *and* dissipative behaviour is mandatory. The nonlinear differential equations that we have introduced containing both conservative and dissipative ranges, seem to fall exactly into Prigogine's formalism. We are tempted to conclude that the "gaps" in the temperature spectrum, and the preferred states of firing of cortical neurons are evidence for such a view.

We look forward to developing this point of view, perhaps with a more rigorous foundation involving other than what are essentially the first moments of probability distributions of the transfer characteristics of neurons, and to keeping continued contact with experimental data.

[1]We are well aware that this differential equation turns up in both population genetics and in epidemiology, and that many of the results obtained in these fields are closely related to our own discussions, as indeed is the application of statistical mechanics to demography and to biochemical rhythms (see [106], [6], [82], [83], [84], and [51]).

8. Acknowledgements

The results noted in this paper were obtained while the author was a member of the Neurophysiology Group of the Research Laboratory of Electronics, Massachusetts Institute of Technology, and was supported in part by the U.S. Army Signals Corps, the Air Force Office of Scientific Research, and the Office of Naval Research; and in part by the U.S. Air Force Aeronautical Systems Division AF 33(616)-7783.
The advice and encouragement of many members of the Research Laboratory of Electronics, especially of members of the Neurophysiology and Communications Biophysics Groups is gratefully acknowledged. In particular Dr. W.S. McCulloch, and Profs. P. Elias and W. Rosenblith helped the author in many ways, as did his friend and colleague S. Winograd.

A. Redundant Automata as Models of Neuron Assemblies

Extended Author's Summary
by
Jack D. Cowan

Electrical Engineering Department,
Imperial College of Science and Technology,
London, S.W.7.

Published in
R. W. Gerard, and J. W. Duyff: *Information processing in the nervous system*
Excerpta Medica Foundation, Amsterdam, New York

1964

A.1. Description of Problem

1. The problem has already been noted during this symposium, Anokhin (see Proceedings of this Symposium), of understanding how the nervous system is integrated. That the nervous system exhibits tremendous complexity of organization is well-known to all students of nervous activity. Some understanding of the principles underlying such organization is a necessary prelude to any understanding of how it all works, if such a "global" understanding is in fact possible.

2. One problem that occurs to the "model-maker" is the following. A necessary prerequisite for the integration of nervous activity is reliability of computation. That is, some reliability must exist in the performance of all the various sub-routines, sequences of operations, and so forth that comprise integrated activity. All available neurological evidence (see for example [154]) indicates that there exist fluctuations in neuro-potentials, large enough to produce Occasional errors or mistakes in neuro-functions. Were such errors to be compounded, in a large mass of neural tissue, the resultant output would be totally independent of the input, so that the tissue could not execute its requisite function, unless there existed some built-in mechanism for reducing the effects of neuro-potential fluctuations. Similar remarks apply to failures due to lesions or damage. The possible nature of this mechanism is of great interest, and forms the subject of this lecture.

3. The theory of most relevance to this problem is information theory, or more specifically the theory of error-correcting codes and the companion theory of the design of reliable automata from less reliable components or modules. By reliability we mean the following: Assume we have a transmission line, e.g. a model axon, a source of symbols, say one's and zero's, and a receiver. If the receiver can identify the patterns generated by the source, and transmitted down the line, in such a way that the probability of error of correct identification of these patterns is less than some given number δ $(0 \leqslant \delta \leqslant 1)$, then we say that δ is an index of reliability of the communication system (source, transmission line, receiver) in question.
   We distinguish computation from communication as follows. Suppose instead of a transmission line (or equivalently, a relay device) we have some unit or Module that responds uniquely to various input patterns. If given an output symbol or pattern from such a module, we cannot, even in the absence of fluctuations, uniquely specify the stimulating input pattern, then we say that the module has executed a computation, e.g. If a given module receives only the input patterns (111) and (000) and responds uniquely with (1) and (0) respectively, then we say that communication has been executed. If on the other hand the module receives all eight possible patterns of zero-one triplets, and responds say, as shown in Table A.1, then we say that computation has been executed, since given an output sequence (i.e. a 1 or a 0) the input sequence cannot be uniquely specified at the receiver.[1] We call the particular mapping of the input patterns on to the set (0,1) the *function* computed by the given unit.

[1]This particular use of the term computation should not be confused with the information reduction mentioned by von Foerster (see Proceedings of this Symposium). When a complete system is engaged in an exchange of 'something' with-its environment that effects a structural change in the system, we have to be very careful in specifying what we mean by information. See MacKay (this Symposium, in discussion following).

Functional description of a computing module	
Input	Output
111	1
110	1
101	1
100	0
011	1
010	0
001	0
000	0

Table A.1.: Possible patterns of zero-one triplets

The problem of ensuring reliable computation with less reliable components thus becomes the problem of ensuring that the mapping by a network of input patterns on to output patterns corresponds to the specified network function, despite fluctuations in the mappings executed by individual modules in the network, i.e. how can high reliability (low δ) be achieved, of a network comprising many complex modules, when it cannot be guaranteed that the reliability of any of these modules is greater than some given δ, ($\delta \geqslant \delta^x$)? Minsky (see the Proceedings of Symposium) in his lecture on the organization of reliable memories, noted that one mode of achieving such a result is to repeat a given computation (or a given memory trace), many times. The probability of error (δ^x) associated with a given network (or information store) is then a monotonically decreasing function of-the number of repetitions. This is an application of a technique developed by the late J. von Neumann [109], in a famous series of lectures. These techniques, while providing sufficient solutions to the stated problem, nevertheless raised some rather difficult questions and problems. From the engineers' viewpoint these solutions were unsatisfactory in that the number of repetitions required to achieve sufficiently small δ^x was always greater than 1,000, or else required infallible modules at many parts of the resultant networks. e.g. If $\delta = 0.005$ for a single module that computes a function of two inputs, then, 30,000 repetitions are required to obtain a $\delta^x = 1.10^{-6}$, and 60,000 to obtain a $\delta^x = 1.10^{-23}$ for the network. This result naturally did not interest engineers (other than the mathematically minded ones) nor did it interest biologists (with similar exceptions).
The question arose as to whether or not some better solutions existed, especially in view of the results provided by a very important theorem in information theory, already mentioned by Professor Brazier (see the Pro-

ceedings of this Symposium) namely the so called "noisy channel coding" theorem ([139], [140]). Mackay (see Proceedings of this Symposium) has already given us an outline of information theory in which the communication system model was described. This model comprised an encoder, i.e. a device that transformed one set of symbol patterns into another set according to certain rules called a *code*, a communication *channel* through which signals propagated, suffering in the process some distortion ("noise"), and a decoder whose function was to recover from the distorted signal patterns ensuing from the channel, the original symbol patterns or *messages*. It was assumed that compared with the channel, both encoder and decoder are noiseless.
Shannon's theorem indicated that if sufficiently long message sequences were encoded into still longer signal sequences via some code, then the probability of error in identification at the decoder of given message sequences could be made arbitrarily small provided certain minimal constraints existed in the encoded signal sequences. These constraints were related to the channel distortion in that the noisier the channel, the greater the minimal constraints before error-free communications could be attained. More exactly, if we call the ratio of the number of signal symbols to that of message signals the *redundancy* of the signal sequences, and its reciprocal, *the rate of transfer of information*, then the theorem expressed the fact that at some critical value of this rate, given infinitely long sequences of symbols, the probability of identification error changed discontinuously from 0 to 1, or rather that there existed at least one code which realised this behaviour. This theorem is to be compared with von Neumann's theorem (equating number of repetitions to redundancy) wherein the error probability is a continuous function of redundancy. The important point to note is that in von Neumann's theorem δ^x approaches zero only with the rate R, whereas in Shannon's theorem δ^x may approach zero provided R is less than some critical value, the well-known channel capacity C. –Thus much greater efficiency (lower redundancies) are exhibited by codes that realise the behaviour embodied in Shannon's theorem.

4. The Question that naturally arises concerns the possibility of extending Shannon's theorem to cover computation as well as communication, i.e. can the theorem be extended to treat, say, an automaton comprising many computing modules, each receiving a large number of inputs from other modules or from an external source, each computing a function of these variables, and each transmitting the resultant output symbol to other modules or to some set of effectors. It has in fact recently been shown (Cowan & Winograd, [170]) that such an extension is possible provided one replaces the entire system of noiseless encoder, communication channel and noiseless decoder by a system or automaton comprising many modules, such that instead of one sub-system performing encoding, another transmission, and

yet another decoding, every component performs decoding, computing, and encoding, in such a way that any one assembly in the automaton will correct errors in its inputs, execute a specified computation, and then encode the results of this computation for transmission elsewhere.
This result has certain interesting implications. Viz: As we have already noted, a code is merely a rule or set of rules for replacing one set of signs by another set. A very simple code would be that which replaces (1,0) by (111,000). If these encoded patterns were transmitted through a noisy communication channel, then (111) might be distorted into any one of the set (111, 110, 101, 011) and (000) into any one of the set (000, 001, 010, 100). The decoder is programmed to replace any member of these two sets by (111) and (000) respectively. In the case of computation however, it turns out to be the case that if we encode the inputs to a computation scheme, and decode the outputs, then we obtain von Neumann's results, or something very like them ([38], [169]). What is required is that the output of the computation scheme, i.e. the value of $f(x)$, not of x has to be preserved, and hence the values of the functions themselves must be encoded, rather than their arguments; i.e. one has to replace a set of functions computed in a specified machine by another set of functions containing constraints and dependencies specified by some code. For example, a simple set of independent parallel computing networks (each with the same topology) which if they were noiseless would serve to compute some specified function, has to be replaced by a larger set of interdependent computing networks that have the ability to cope with noise, in the way that we have already outlined.
Figure C.21 shows an automaton consisting of four independent parts, e.g. a parallel neuronal network wherein each unit produces a specific response to certain input patterns. Figure C.22 shows an automaton in which the input rank computes and encodes, the second rank decodes the perturbed output of the first rank, computes and encodes, and the final rank decodes and computes. Errors occurring in the first and second ranks are corrected, but not of course errors in the final output. However it is enough that no intermediate errors are allowed to propagate. Under certain circumstances, the probability of error δ^x of this automaton is less than the probability of error δ of the precursor shown in Figure C.21. It will be seen that the redundant modules of the coded automaton are more complex than the modules of the precursor. This results from the fact that decoding, computing and encoding operations have all to be performed by the complex modules and it is assumed that the error probability associated with the complex module is not greater than the error probability associated with the simpler units of the precursor. It is also of interest to note that if errors in the pattern of interconnexions occur, then these too may be controlled by the introduction of suitable redundancy (Cowan & Winograd, op.cit.) We have thus outlined, albeit in a somewhat cursory manner, methods for the synthesis of redundant

automata that compute reliably, complex functions, despite errors both of function and structure.

5. It is of interest to note that certain necessary requirements exist on the logical structure of these automata. In particular, there must exist both excitatory and inhibitory connexions to all units, on the average in about equal numbers; there must also exist a large number of presynaptic inhibitory connexions ([102], [15]) to each unit; and since there must exist a large number of overlapping connexions, there must exist by implication an appreciable large number of overlapping inhibitory connexions, i.e. lateral inhibition is an important feature of these networks. It is also worth noting that the overall functional organization displays a multiple diversity, i.e. any one module computes a mixture of many precursive functions, and any one function is stored in many places, rather similar to that displayed by the models constructed by Beurle, [12] that were based on an axiomatization of the known structural organization of neural tissue.

A.2. Acknowledgement

The results noted in this paper were obtained while the author was a member of the Neurophysiology Group of the Research Laboratory of Electronics, Massachusetts Institute of Technology, and was supported in part by the U.S. Army Signals Corps, the Air Force Office of Scientific Research, and the Office of Naval Research; and in part by the U.S. Air Force Aeronautical Systems Division AF 33(616)-7783.

The advice and encouragement of many members of the Research Laboratory of Electronics, especially of members of the Neurophysiology and Communications Biophysics Groups is gratefully acknowledged. In particular Dr. W. S. McCulloch, and Profs. P. Elias and W. Rosenblith helped the author in many ways, as did his friend and colleague S. Winograd.

B. Neuronal Nets

by
Jack D. Cowan

Electrical Engineering Department,
Imperial College of Science and Technology,
London, S.W.7.

To be Published in ENCYCLOPEDIA OF INFORMATION,
LINGUISTICS AND CONTROL
Pergamon Press 1967.

July 1965

B.1. Introduction

Neuronal nets are aggregates of interacting nerve cells or "neurons". Neurons are the basic structural units of animal nervous systems. To survey in any real detail the experimental background involved is of course completely out of the question. The reader is referred to the standard reference on the subject (see [43]) for authoritative surveys of the subject matter. Here we give only a brief survey. Figure C.23 shows lateral and medial aspects of vertebrate cerebral hemispheres with some of the more prominent features indicated. These features represent constellations of neuronal nets, or "super nets" whose patterns of interconnexion or "wiring-diagrams" are determined on a hereditary basis, in the large if not in the small. Among these super nets perhaps the telencephalon, diencephalon, and

rhinencephalon are of most interest, being concerned with "higher" functions of the nervous system– perception, learning, conceptual thinking and so on. The other super nets, the mesencephalon, metencephalon, myelencephalon and spinal cord, are concerned with "lower" functions-reflex behaviour, autonomic activities, attention and so on. They also serve to relay signals to and from the higher super nets, from and to sensory receptors and muscles. The rhinencephalon, which includes the hippocampus and part of the orbital surface of the frontal lobes (see Figure C.23), can be considered to be the visceral and olfactory brain. It apparently plays an important role in determining the long-term behaviour of animals, since it appears to exert some degree of control over the addressing of signals to the storage centres supposedly located in the telencephalon, [172]. The diencephalon, which includes the neuronal nets of the thalamus and hypothalamus, plays a very important role in the integration of the total activity of the central nervous system (CNS). From thalamic nets develop the retinas and optic nerves for example, and upon such nets converge incoming or "afferent" paths conveying signals from receptors. It acts as a relay station where afferent signals are processed and sent to the telencephalon. There is also diffusion and spread of signals to other super nets.
What appears to be a very general rule in the anatomy of the nervous system is found in diencephalic interconnexions with the telencephalon, namely that there are two-way interconnexions between nets. For example the Medial Geniculate connects, with the superior temporal gyrus and vice-versa, the Pulvinar connects with parietal and temporal areas of telencephalic cortex and vice-versa. At a lower level there are two-way interconnexions between diencephalic nets and non-specific nets located in the myelencephalon – the reticular formation. This two of three tiered arrangement of neuronal nets is of great importance.
The telencephalon is undoubtedly the most important of the super nets of the nervous system with respect to the execution of higher functions. The Cerebral Cortex is that part of this super net which is especially concerned with such functions. The microstructure of this cortex is heterogeneous and contains a large number of cell types and wiring-diagrams. Grossly speaking there are two types of tissue to be seen – "grey matter" comprising the "cell bodies" and "dendrites" of neurons and "white matter" comprising the "axons" or outgoing processes of such neurons. The neurons are arranged in fairly definite layers. There are about five different cell types of which the most important are the "pyramidal" cells and the "stellate" cells. Axonal bundles enter the cortex from subjacent white matter generally at a direction normal to the cortical surface. In this way cortical nets are separated into columns normal to its surface. Many of the axon bundles comprising the white matter are corticifugal, running as "association" fibres to other nets in the ipsilateral hemisphere, or as "commisural" fibres to corresponding nets in the contra-lateral-hemisphere, or as "projection" fibres to extra-cortical nets. Similarly there are corticipetal fibres from such nets and association and commisural fibres from other cortical nets. Figure C.24 shows some of these details.
The cerebral cortex does not have a uniform histological structure. There are

differences in the distribution of cell types, their number and packing-density. Indeed there is a highly specific selection of cell types and wiring diagrams, at least in the large, with laminar and columnar organization very common. Important cortical nets are the "Projection" cortices – visual, auditory, and somaesthetic, which receive afferent stimuli from peripheral nets, and the motor cortex which transmits signals to peripheral effectors. As far as the microstructure of such nets is concerned it seems that averages are the only stable histological parameters. D.A. Sholl [142] has estimated many of these parameters, the more important of which are the number and packing-density of neurons, the mean size and shape of the neuronal membrane – the "perikaryon", the length and distribution of axons, and the origin and distribution of afferent fibres. As a result of Sholl's calculations it may be estimated that as much as $0.1\,\mathrm{mm}^3$ of cortex has its state modified as an immediate result of the activity in a single afferent neuron.
This then is the anatomical substratum upon which physiological aspects must be superimposed. It is evident that the laminar and columnar organizations of cortical nets, together with the tiered arrangement of super nets coupled everywhere by two-way interconnexions, makes for a very complex system within which the coding, addressing, transmission, storage and retrieval of signals takes place. The investigation of such processes is one of the aims of neurophysiology.

B.2. The Neuron

Neurons are cells usually not more than 0.1 mm in diameter, which have become specially adapted to processing and transmitting electrical signals within the nervous system. Each neuron comprises a "soma", together with dendrites, and a single axon, (see Figure C.25). The dendritic complex functions as an antenna, collecting incoming signals and conveying them to the soma while the axon functions as a transmitter, conveying outgoing signals to other neurons and to muscles. Such signals generally have to cross separating gaps between neurons, called "synapses". How they do so has been the subject of much recent work ([36]). The neuron and its processes are separated by a lipid-protein membrane from the surrounding extracellular medium. The somato-dendritic or "SD-membrane" of the neuron is very sensitive to incoming signals. In recent years direct studies of the nature of this sensitivity have been carried out via intracellular and extracellular recording by microelectrodes of very small tip diameter, about 1 µm. It has been established that the SD-membrane has a total capacitance of about 3 nF and a total conductance of about 1 µS. The membrane as a whole responds to incoming signals effectively like a parallel RC-circuit with a time constant of about 3 ms. A very important factor in such a response is the influence of varying conditions at the synapses. Electron micrographs show that the neuronal membrane surface is covered by many button-like processes or "Boutons", which have been identified as the regions of interaction with other neurons. The gaps between such boutons and the cell

membrane are the synapses. The physicochemical phenomena involved in synaptic transmission are of great complexity and interest. Roughly speaking, electrical potentials arriving at some synapses induce "presynaptic potentials" which liberate certain chemicals or "transmitter substances", which appear to migrate across the synapses, causing potential changes in sub-synaptic regions of the SD-membrane, called "postsynaptic potentials". The equilibrium or "resting potential" of the membrane is about $-70\,\mathrm{mV}$ (with respect to some reference electrode), and the postsynaptic potentials act either to depolarize the membrane towards $0\,\mathrm{mV}$, or to hyperpolarize it towards $-80\,\mathrm{mV}$. When the membrane is sufficiently depolarized a breakdown in its resistance occurs, resulting in a rapid swing in the membrane potential leading to the production of an outgoing impulse from the neuron. Depolarizing postsynaptic potentials are known as excitatory-postsynaptic potentials or EPSPs and hyperpolarizing ones as inhibitory postsynaptic potentials or IPSPs, since they (respectively) contribute to or detract from the state of excitation built up within the neuronal membrane. Synapses can be either excitatory or else inhibitory. Depolarizing Synapses can be either excitatory or else inhibitory.
An equivalent electrical circuit has been inferred for the SD-membrane and associated synaptic regions of neurons in the output or motor side of the nerve trunks in the spinal cord. It should be realized that there are different neuronal types which have varying modes of response to stimuli, and that the output or "motoneuron" equivalent circuit is not necessarily typical of neuronal responses within the nets of the CNS proper, as distinct from the spinal cord. In particular, dendritic effects certainly influence the way in which excitation is built up within the cell membrane and probably play an important role in CNS function. The motoneuron equivalent circuit is shown in Figure C.26 ([35]). G_e and G_i are variable conductances composed of hundreds or even thousands of individual synaptic conductances such that the highest observed value for G_e is of the order of $2\,\mu\mathrm{S}$, that for G_i slightly less. The membrane conductance and capacitance are as previously noted. The resting potential of the membrane is represented by the battery generating a voltage $\mathscr{E}_m$ of $-70\,\mathrm{mV}$, and the potential driving the IPSPs by the battery generating a voltage $\mathscr{E}_i$ of $-80\,\mathrm{mV}$.
Consider first the effects of a single excitatory and a single inhibitory synapse of conductances g_e and g_i respectively, connected to the membrane circuit at time $t = 0$. Following excitation by presynaptic potentials, transmitter substances are released which effectively result in a short-circuiting of sub-synaptic SD-membrane regions, so that currents flow which either charge or discharge through the membrane, depending upon the nature of the transmitter substance. The chemical nature of transmitter substances, especially that serving inhibition is still an open question. It is easily shown that the response is given by the equation

$$e(t) = \mathscr{E}_m + \frac{g_i}{g'} \cdot (\mathscr{E}_i - \mathscr{E}_m) \cdot \left(1 - e^{\frac{g't}{C}}\right) - \frac{g_e}{g'} \cdot \mathscr{E}_m \cdot \left(1 - e^{-\frac{g't}{C}}\right) \qquad \text{(B.1)}$$

where g' is the total circuit conductance equal to $g_m + g_e + g_i$. After times t_e and t_i both of the order of 1 ms, the presynaptic excitations cease, and the circuit returns to its equilibrium value with a time constant of $g_m^{-1}C$ secs. If the excitation is solely inhibitory or else excitatory, either g_e or else g_i is zero, and the individual responses are readily obtained. Figure C.27 shows some typical responses of the system.

We note that the calculated EPSPs and IPSPs only approximate to those observed in vivo. Firstly there is a delay of about 0.5 ms between the onset of a presynaptic potential and the build-up of a postsynaptic potential. Secondly observed EPSPs usually have time constants longer than 3 ms. It has been suggested that these effects are due to the finite times taken to build up and to remove concentrations of transmitter substance ([36]). It has also been suggested that the effects are due to the influence of the dendritic complex, non-negligible even in motoneurons, which in effect means that the "lumped" equivalent circuit of Figure C.26 must be replaced by a "distributed" circuit ([124]). The resulting electrotonic effects tend to smear out EPSPs and IPSPs induced in peripheral regions of the dendritic complex so that their form at the soma is somewhat distorted, the onset and rate of decay being slowed down. Figure C.28 shows an example of the results of Rall's calculations.

As we have noted the SD-membrane is covered with synapses, so it is continually being charged and discharged. Every change ∂g_i in the total inhibitory synaptic conductance generates a current $(\mathscr{E}_i - \mathscr{E}_m)\,\partial g_i$ mA (in the lumped parameter model) charging the membrane, and every change ∂g_e in the total excitatory synaptic conductance generates a current $-\mathscr{E}_m\,\partial g_e$ mA discharging the membrane. When the membrane potential reaches a certain minimum level, the conductance g_m changes suddenly and the result is a rapid change of potential, followed by a slower recovery to the resting potential (usually after a prolonged period of hyperpolarization), when presynaptic potentials again begin to affect the membrane. The phenomenon is complicated in that the sudden potential change, known as the "Action Potential" has several components, and is generated at specific sites in the neuronal membrane. In motoneurons it is the initial segment of the axonal membrane – the "IS-membrane", which initiates the action potential. Potential changes in the IS-membrane, which has a lower threshold than the SD-membrane for resistive breakdown, elicit even larger changes in the SD-membrane and wipe out the existing potentials there to produce the fully-fledged action potential. We therefore consider axonal membrane, particularly the IS-membrane. Its equivalent circuit is simply the right-hand side of the circuit shown in Figure C.26. To account for the breakdown of the membrane resistance, the underlying ionic phenomena have to be considered. The basic idea is embodied in the sodium potassium ion pump postulated by A. L. Hodgkin and A.F. Huxley, [69]. In the resting state the membrane permeability to potassium ions is supposedly much greater than that to sodium ions. When the membrane potential is sufficiently large, an increased permeability to sodium ions is created which, if sufficiently greater

than the potassium ion permeability, causes a change of membrane potential from the negative potassium potential to the positive sodium potential. Thereafter a recovery period occurs when the membrane returns slowly to its resting potential. An equivalent circuit which accounts for the ionic changes is shown in Figure C.29. $\mathscr{E}_k$, the diffusion potential for potassium ions, is given by the Nernst equation

$$\mathscr{E}_k = 58 \cdot \ln \frac{[K]_0}{[K]_i} \tag{B.2}$$

where $[K]_0$ and $[K]_i$ are respectively the outside and inside concentrations of potassium at the membrane. A similar equation governs the flow of sodium ions and other species. The resulting membrane current is given by the equation

$$i_m = C \frac{\mathrm{d}e}{\mathrm{d}t} + m^3 h g_{\mathrm{Na}} \cdot (e - \mathscr{E}_{\mathrm{Na}}) + n^4 g_k \cdot (e - \mathscr{E}_{\mathrm{K}}) \tag{B.3}$$

where m equals Na activation, $h-$ Natrium inactivation, and $n-$ Kalium activation. These variables are dimensionless, varying between 0 and 1, each satisfying an ordinary differential equation involving $e(t)$, $\mathscr{E}_m$ and the absolute temperature T. The Hodgkin-Huxley model is thus based on four ordinary non-linear differential equations. No general solutions of these equations are known but they have been solved by numerical integration and by analogue Computer simulation. R. Fitzhugh, [44], obtained the solutions shown in Figure C.30 using such a computer. The most notable feature of the solutions is the existence of a threshold for the production of the fully-fledged action potential or "spike". Such spikes are produced as a result of the excitation built-up in the neuronal membrane and are transmitted by the axon to other cells, propagating without decrement until they reach the terminal arborization of the axon. Because there is a threshold for spike emission, and because the spike propagates without decrement with an amplitude depending only upon local conditions in the axon and independent of stimulus strength, the process was called "All-or-None" emission ([3]).
Although there is no direct relationship between spike parameters and the intensities of stimuli acting on the SD-membrane, there is one between the mean rate of spike emission and the stimulus intensity. A number of experimental determinations have been made of this relation ([9], [56], [55], [154], [81], [80], [68], and [112]). Interpretation of the experiments is difficult in that data are taken from widely varying preparations (Frog, Rat, Cat; isolated axon, motoneuron, pyramidal neuron; constant stimulating currents, intermittent current pulses), and various secondary effects such as adaptation and accommodation of the SD-membrane to constant stimuli play a role in determining the shape of the rate-intensity curves obtained. Roughly speaking the mean rate of spike-emission is a monotonic increasing function of the quantity i/i_0, where i is the intensity of the applied current, and i_0 is the "rheobase", the smallest current which will elicit a spike. In general there is a maximum rate for spike-emission. This occurs because the membrane requires

some time to recover its sensitivity after emitting a spike, and for some time after the emission of a spike the membrane is hyperpolarized. The minimum interspike interval is called the "refractory period". Figure C.31 shows three examples of rate-intensity curves, one linear, one exponential, and one sigmoidal. It seems to be a general finding over a wide range of stimulus intensities, that if the reciprocal of the interval between the first two spikes emitted by a cell is taken as the firing-rate, then the rate-intensity function is sigmoidal. There is indirect evidence supporting such a conclusion in recent observations of correlations between Post-Stimulus Time Histograms (PSTH) of spikes emitted by cortical neurons, and the changes of membrane potential recorded on a common microelectrode ([128], [27]).
All these rate-intensity functions represent averages, and there is much variability in the emission of a single spike. The source of this variability may be either entirely outside the cell, e.g. receptor potential fluctuations ([98]), or in synaptic or neuro-muscular junctions ([92]; [40]; [2]), or there may be intrinsic fluctuations in the neuronal membrane itself ([14], [112], [63], [155], and [156]). Recently Verveen and H.E.Derksen ([157], [158]) have shown that the spectrum of the potential fluctuations in axonal membrane is predominantly of the $1/f$ type found in current-carrying carbon resistors, thin metal films, and semiconductor devices. These particular fluctuations are related to the flux of potassium ions across the membrane. The extension of these results to SD-membrane let alone to neurons in vivo is an open problem. However there is evidence of variability in the many studies which have been made of the spikes emitted by "spontaneously" active neurons, and of "driven" ones in various nets within the CNS, (see for example [133], [122], [160], [130], [90], [44], and [8]). All these experiments lead to the conclusion that there are numerous sources of variability acting upon the neuronal membrane, some of which are linear in their effect, some nonlinear. A statistical description is ultimately required of the neuronal response. The existence of sigmoidal and exponential rate-intensity curves evidently implies the existence of unimodal probability-density functions of spike-emission.
As we have already noted these rate-intensity curves cover only the first few milliseconds of the neuronal response. The main reason for this is that the neuron "adapts" to constant stimuli so as to limit the number of spikes it emits. Similarly the neuron "accommodates" to stimuli in that a gradual increase in stimulus intensity eventually fails to evoke a spike discharge. The time constants of these processes are of the order of tens of milliseconds. Both processes are thought to be due to a slow increase in the threshold for spike emission, brought about by a reaction in the membrane opposing the applied current, tending to re-establish equilibrium.
Many of the concepts we have discussed have been derived from work on motoneurons. For other neuronal types it is certainly true that there are significant departures from motoneuron-like behaviour which are important features of signal processes. We have already cited Rall's work on the dendritic complex. We reiterate that local potential changes in sub-synaptic regions of SD-membrane spread electrotonically to the IS-membrane. Clearly the total excitation built-up there may

be weighted by a spatial as well as a temporal factor, so that dendritic geometry may well play a very important role in information processing within the CNS. In addition, because of differences in the neuronal membrane, different regions may be in different dielectric states so that current flows continually between them. In peripheral systems dealing with sensory transmission and with the regulation and control of such vital processes as respiration, it appears that many neurons have a "gain" greater than one, i.e. they emit more spikes than they receive ([171]), an effect which is sometimes ascribed to oscillations in the after-potentials which follow spike-emission.
With these reservations in mind we conclude this section having built-up a picture of the neuron as a cell which converts neuronal "messages" represented within spike patterns into continuous signals, by processes of spatial and temporal summation and integration, and then sends its own "to whom it may concern" message within its output spike pattern. If, as some neurophysiologists contend, the neuronal message is partly coded in the form of variation of the mean spike-rate or equivalently of the mean interspike interval, then it will be seen that the neuron acts essentially as a pulse-interval modulator with spike trains of some mean rate serving to carry signals.

B.3. Neuronal Nets

So much for the individual neuron. We now turn to neuronal nets proper. We shall look at only three aspects of nets – the responses of single neurons embedded in a net, the responses of nets inferred from recordings of the activity of large numbers of neurons, known as "electro corticograms" or ECoG when such recordings are made from the exposed surface of the brain, and the responses of immense numbers of neurons, inferred from recordings of electrical activity through the scalp, the "electroencephalogram" or EEG.

B.3.1. The Electroencephalogram

When gross electrodes (about 1 cm in diameter) are placed on the surface of the scalp and the activity measured either monopolarly or else bipolarly, a fluctuating voltage (about 100 µV) can be seen. Because of the impedance characteristics of the System and recording apparatus, the recorded voltages are a measure of the electrical activity of supernets seen through a low-pass filter. Thus the EEG is a record of the activity essentially under a monopolar electrode or else it is the algebraic sum of the activities, under bipolar electrodes. Its importance is that it serves as a relatively stable measure of physiological significance.
The EEG as a whole is a complex comprising many varying patterns, and is characterized by such statistics as its autocorrelogram, power spectrum, amplitude distribution and envelope amplitude distribution. Figure C.32 shows a typical sam-

ple of EEG from a normal adult. Figure C.33 and Figure C.34 show a correlogram and spectrum, and amplitude and envelope amplitude distributions respectively, of typical samples of EEG. It will be seen that there are prominent rhythmic components in some of the samples. These were discovered by H. Berger, [11], who described the EEG as essentially two basic rhythms, the alpha and beta rhythms. The alpha rhythm is usually in the 8-13 cycles per second range when it exists, is prominent in parieto-occipital regions, and disappears in a visually attentive subject. However a well-developed alpha rhythm is a relatively uncommon finding in a normal adult, most samples showing instead low voltage fast activity in the frequency band. The beta rhythm is also a low voltage fast activity between 18 and 30 cycles per second, and is prominent in precentral and postcentral gyri. It will be seen from Figure C.34 that the EEG spectrum is indeed concentrated in the alpha and beta bands. The autocorrelograms do not imply that there is a stable oscillator generating these rhythms, and in fact, narrow-band Gaussian noise gives autocorrelograms and is indistinguishable from EEG ([163]). In addition the amplitude and envelope amplitude distributions are approximately Gaussian and Rayleigh distributed ([163], [137]). These properties are of course also found in narrow band Gaussian noise.
As we shall discuss later, there is evidence that EEG rhythms are generated in the supernets of the diencephalon and telencephalon under control by myelencephalic signals. Recent work by D.O. Walter et.al., [161], suggests that there are at least two different generative processes at work – a process which spreads homolateral from posterior occipital regions, generating significantly coherent activity in the same band. The identification of these processes with diencephalic nuclei has not been made. It has been suggested that the beta rhythms are harmonically related to the alpha rhythms or that they are the result of a difference in myelencephalic control signals ([18]). It is fair to say that the detailed mechanisms underlying the generation of the EEG remain to be discovered. In spite of this lack of understanding, the EEG has proved to be of value as an index of "physiological state", and an important aid to the diagnosis of pathological conditions in the brain.
Figure C.35 shows the sequence of changes in the normal EEG during alertness, drowsiness, and sleep. It will be seen that there is a progressive change from the low voltage fast activity of the alerted state, through successive states characterized by higher voltages and lower frequencies, to the "delta" waves of deep sleep. It will be seen that the waves in all these records are roughly symmetric and could conceivably be generated by linear oscillators. This is not the case with the EEGs typical of pathological conditions such as Epilepsy and other diseases. Figure C.36 shows typical patterns found in "Petit-Mal" and "Grand-Mal" Epilepsy, the former showing the well-known "Spike-and-Wave" phenomenon, the latter characteristic high-frequency spike activity. In general terms it might be said that the frequencies found in the analysis of the EEG are the most obvious index of physiological state, although the shape and amplitude of the activity is also of some interest. Figure C.37 shows a schematic representation of the frequency spectra associated

with different physiological states.

B.3.2. Electrocorticograms and Evoked Responses

So far we have considered only the spontaneous activity of the nervous system. A very important source of knowledge concerning the system is its response to stimuli. Such response may be studied by means of the gross surface electrodes used in electroencephalography, or else smaller extracellular electrodes may be inserted into selected regions of the system to sample the local responses of neuronal nets. In either case, the recorded evoked potentials differ from the spontaneous activity in that they have a definite relationship to the onset of the stimulus, a constant pattern of response and a focus of maximal response in the system. Stimuli may be natural, or else clicks, flashes, and so on, or else stimulating electrodes may be inserted in other neuronal nets to produce signals which in turn stimulate the observed net.
Figure C.38 shows a fairly typical record ([16]) of an evoked potential in the visual cortex of a lightly anaesthetized cat responding to a light flash of 10 µs duration. The record is an average over fifty flashes given at 1 ms intervals. According to convention the downward deflection corresponds to a positive surface potential, the upward one to a negative potential. The biphasic nature of the response a surface-positive deflection followed by a longer lasting surface-negative deflection, is typical of most evoked responses. In addition the surface-negative deflection is often followed by long-lasting after-effects such as rhythmic activity. According to volume conductor theory the surface-positive deflection is due to signals from the periphery approaching the cortex-presynaptic excitation, whereas the surface-negative deflection is due to signals within the visual cortex more-or-less receding from the recording electrode – most probably postsynaptic excitation induced in the dendrites of cortical neurons by presynaptic excitation. Figure C.39 shows another typical evoked response ([17]) to flash in Man. Here there is a noticeable after-discharge phase-locked to the flash. Such a record is quite common, the rhythmic after-discharges evidently having some significant functional role. Similar evoked responses are found in all those areas of the cortex which first receive signals from the periphery – the projection areas. There appear to be topological maps from receptor areas through intermediate diencephalic nuclei to these projection areas. So neighbourhood relations on "receptive-fields" are preserved to the projection areas, and the evoked response is a local response, a change of state in perhaps 1 mm^3 of cortex caused by stimulation of some particular set of receptors. However the receipt of signals from such sources is not restricted to the primary projection areas of the cortex. For example, in somatic-sensory cortex, second and even third receiving areas have been found. Similar considerations apply to the visual and auditory cortices. Finally in some areas the initial biphasic evoked response is followed by a larger and longer-lasting biphasic response, the "secondary" response. This is much less localized than the primary evoked response,

and it is established that it reaches the cortex by a different path ([33]). The primary response represents activity in thalamo-cortical fibres resulting from the stimulation of thalamic nuclei by signals from peripheral nets. These signals follow the "classical sensory pathway" to the contralateral cortex via the lateral division of the medial lemniscus and the ventrolateral nuclei of the thalamus. The secondary response is due to peripheral signals following a different path which does not traverse the thalamus.
In addition to the responses evoked by stimulation of receptive fields, there are those evoked by stimulating electrodes placed in different regions of the brain. These responses can be either specific, local responses to the stimulus, or else non-specific generalized responses of supernets appearing as changes in the EEG. In either case they provide valuable information into the nature of the interactions between supernets, particularly the mechanisms responsible for attention. For example G. Moruzzi and H. W. Magoun, [107], discovered that maintenance of the waking state does not depend on inflow along the classical sensory pathway, which would give rise to specific evoked responses, but instead depends on signals in the "ascending reticular system" comprising nets in both the diencephalon and myelencephalon. Stimulation of this system produces changes in the EEG from slow to fast activity, sometimes over the whole cortex, sometimes only in fronto-parietal nets. Sensory modality is apparently signalled by such an influx. In contrast to this the hippocampus changes from fast to slow activity, as seen in the ECoG ([57]). Dempsey and R.S. Morison, [34], found that slowly repeated stimulation of medial nonspecific nuclei in intra-laminar nets of the thalamus at rates not too different from that of the animals own cortical potentials can entrain these rhythms to the stimuli. Such responses are called "recruiting" responses. These differ from the augmenting responses seen on stimulation of other nuclei. Recruiting responses are of long latency, widely distributed over cortical nets, and are predominantly surface-negative. They have the characteristics of dendritic potentials initiated in the upper cortical layers. This surface-negative component is depressed by simultaneous high frequency stimulation of the reticular system. Augmenting responses are local responses of short latency restricted to the projection areas served by the thalamic net which is stimulated, and they are biphasic with an initial surface positive component. It will be seen that differences exist in evoked potential and ECoG records which are related to the specificity of the nets involved. It seems that alpha and beta rhythms, recruiting responses, and secondary components in the evoked response result from the interaction of nets within the upper five cortical layers with nets of the thalamic myelencephalic and mesencephalic reticular systems. The primary components of the evoked response together with augmenting responses appear to result from the interaction of nets in middle cortical layers with nets in the ventro-lateral nuclei of the thalamus.
Broadly speaking we may assume that there are essentially two different subsets of supersets within the nervous system, one dealing with the classification and encoding of specified stimulus attributes, the other with holistic aspects of the CNS

in relation to its environment, i.e. with the general physiological state of the animal. We may term the former aspect "phasic", the latter "tonic". There is anatomical evidence for the specificity attributed to phasic responses in the topological maps from receptive-fields to projection areas. It has been suggested (Dempsey and Morison op.cit.) that the "nonspecific fibres" ([123]) of unknown origin which terminate in the upper layers of the cortex, may be efferents from reticular nets especially thalamic ones. Brazier (op.cit.) has noted that these fibres could synapse with the dendrites of pyramidal cells in the cortex, and has suggested that such an interaction could give rise to tonic activity. There have been many suggestions that cortical potentials measured by gross and macroelectrodes represent both presynaptic and postsynaptic excitation in the dendrites of pyramidal neurons. The surface-positive components of the evoked response could well be the sum of presynaptic potentials and the surface-negative components together with the waves of the EEG are equally likely to be summed EPSPs and IPSPs, seen through some kind of low-pass filter. Recent work by M. R. Klee et al.,[86] has shown that there are significant correlations between waves in the EEG and slow membrane transients in single neurons in cortical nets during the course of spontaneous activity and also during augmenting and recruiting responses. Some light is also shed on this possibility by measurements of the changes in D.C. potentials on the cortical surface (measured with respect to a reference electrode usually placed in a ventricle) associated with evoked potentials, (see [110]). H. Caspers, [24], has suggested that a "dynamic balance" exists between EPSPs and IPSPs in dendritic regions of cortical pyramids, which is reflected in both the D.C. potential and the evoked responses. The "resting" level is supposed to be a state in which on the average, depolarizing effects cancel with hyperpolarizing ones. A positive or negative shift of the D.C. potential increases the average number of units which will respond to an input by a relative increase in either depolarization or hyperpolarization.
Caspers concludes that the evoked potentials may represent a modulation of the D.C. potential. It is of interest that W. Grey Walter, [162], has found changes in the EEG – "E-waves" – with time-constants of about 1 s which appear to relate the state of human subjects to stimulus contingencies. E-waves appear to be closely related to the cortical D.C. potentials.
Direct stimulation of the cortical surface (Eccles 1951) has led to the suggestion that the observed superficial response, a brief surface-negative wave, corresponds to the focal record of EPSPs induced in the "Apical" dendrites of pyramidal cells by presynaptic impulses. Experiments on the changes accompanying spreading depression ([94], [52]), and on isolated slabs of undercut cortex ([21]) relating changes in A.C. potentials and evoked potentials caused by direct cortical stimulation to changes in the firing patterns of individual neurons, all support the conclusion that postsynaptic potentials in dendrites constitute the main component of the EEG and of the evoked responses.
These conclusions are based on experiments with corticipetal fibres. As we have previously noted it seems that for every corticipetal fibre there is a corticifugal

one. So we have to consider the possibility of physiological feedback from cortical nets to thalamic, mesencephalic, and myelencephalic ones, and the role it serves in maintaining or modifying the activity of the CNS. Evidence is available concerning corticifugal pathways playing upon the mesencephalic reticular formation. G. F. Rossi and A. Zanchetti, [136], have summarized the anatomical evidence. Descending "Suppressor" effects induced by cortical stimulation ([103]) strongly suggest that these cortico-thalamic paths are partly inhibitory in nature, and indeed A. Hughelin et al., [75], have shown that the reticular formation is under the restraining action of the cortex. At diencephalic levels P. Andersen et.al., [5] have demonstrated that stimulation of certain cortical areas causes inhibition in peripheral afferent nets. I. Calma, [23], has demonstrated a cortical influence, mainly inhibitory, on some of the thalamic relay nuclei in the classical sensory pathway – the ventrobasal and posterior thalamic nuclei. In addition to such "vertical" circuits, it appears that there are "lateral" circuits between homologous areas in the two hemispheres. W. P. Koella and A. Ferry, [88], have found that transcortical electrical polarization of one hemisphere interferes with the evoked potentials not only in the polarized ipsilateral cortex, but also in the contralateral one. Within a limited range the changes observed in the contralateral hemisphere were opposite to those in the ipsilateral cortex. Indeed the experiments show that surface-positive polarization decreases the "arousal" level in the ipsilateral cortex and enhances it in the contralateral cortex, while surface-negative polarization has the opposite effect. This suggests that there are negative-feedback circuits between cortical and mesencephalic nets. Koella and Ferry noted that a particular aspect of such a system would be that its output diverges bilaterally to impinge upon two separate effector areas, and that the feedbacks from these two areas converge upon the brainstem and operate through a mixer to exert their combined restraining influence. Figure C.40 shows some general features of the postulated system. Because of the mixer (which is just a summing device) such influences can be excitatory or inhibitory. This suggestion is of interest in the light of the findings of Bremer, [19], [20], H. T. Chang, [25] and C. A. Marsan and A. Morillo, [101] of both excitatory and inhibitory influences of commisural fibres. Finally we should add that many cases of physiological feedback within nets have been discovered. A frequently found pattern is for excitatory cells to synapse via "recurrent collaterals" with inhibitory "interneurons" which in turn synapse with the excitors. Such circuits have been described in the cerebral cortex ([115]), in the hippocampus ([77]), in the ventrobasal complex of the thalamus ([4]), and in the lateral geniculate ([138]). Such feedback "oscillators" have in fact been postulated to provide the basis for the slow rhythms which we have discussed (Andersen and Eccles op.cit.). This then concludes our survey of the slower potentials of neuronal nets. We have of course looked at only a small fraction of the relevant experimental data, and we have certainly oversimplified. Nevertheless we can surely conclude that there are broadly speaking, tonic and phasic components in the slow potentials, that both presynaptic and postsynaptic excitation in neuronal nets give rise in some way to

these potentials, and that the multiplicity of circuits between and within nets is crucial for the maintained functioning of the system.

B.3.3. Microelectrode Recordings and Slow Potentials

We now turn to the problem of integrating what we have learned concerning slow potentials with the spike-patterns that can be recorded by extracellular and intracellular microelectrodes. B. Renshaw, A. Forbes, and B. R. Morison, [126], first recorded spikes emitted by neurons within the nets of the cortex, but they could find no correlation between spike patterns and slow potentials recorded from the hippocampal surface. C.-L. Li and H. Jasper, [96], obtained simultaneous records of spike patterns and electrocorticograms (see Figure C.41). Jasper concluded that there was no clear-cut relationship between the two different records. The problem is that there is great variability in the records. The gross electrodes apparently pick-up only coherent activity in a relatively large volume of tissue – about $1\,mm^3$. The microelectrode picks up all the activity in that part of the neuronal membrane which is near the electrode tip. In recent years, developments in the technology of data processing have made it possible to compute statistical measures of the activity sampled by both gross and microelectrodes (see [134]). In this way correlations can be found between spike patterns and slow potentials. In section B.2 we discussed determinations of rate-intensity curves for isolated neurons, and noted that correlations have been found between PSTHs and mean slow potentials, obtained from records taken with a single microelectrode of the activity of a neuron within a cortical net. Figure C.42 shows such records with obvious correlation between them. S. S. Fox and J. H. O Brien, [45], have obtained somewhat similar results, and because the slow potentials are still recorded even after the observed neuron has been killed, have inferred that the slow potential recorded on the common microelectrode is a close approximation to the evoked response of the net containing the observed neuron. The inference drawn from this is that the frequency distribution of the firing of a single neuron corresponds to the mean evoked response of the net. Such a conclusion implies a very high degree of synchrony within such a net, and it is perhaps doubtful that the postulated relation is generally true. However Robertson, [127] has found correlations between the mean firing rate of cortical neurons in anaesthetized preparations, and the level of surface-positive polarization of the cortical surface, which is known to effect the amplitude of the evoked response (Caspers op.cit.). In addition to these results relating to the evoked response, there is also evidence relating to correlation between oscillations in slow potentials and temporal patterns of neuronal firing known as "bursts". In the spinal cord C. Terzuolo and B. E. Gernandt, [150], have found burst activity to be coincident with the rising phase of the slow potential. M. Verzeano and K. Negishi, [159], have found similar correlations in records from cortical and thalamic neurons. Robertson (1966b op.cit.) has also found correlations between neuronal spike bursts and "spindles" in the EEG. Verzeano and Negishi's results

are also of interest in that working with multiple microelectrodes. They found circulating spike activity in just those corticothalamic circuits which are held to be responsible for the tonic components of neuronal activity.
All these different considerations certainly indicated that the slow potentials reflect the postsynaptic effects of afferent input to neuronal nets, and possibly the efferent flow from nets. Experimental results seem to indicate that the mean spike emission rate in neuronal nets is a monotonic function of such parameters of the slow potentials associated with the net, as amplitude and rate of change. E. R. John et al., [76] conditioned animals to respond to periodic stimuli and noted a tendency towards increased correlation between evoked potentials and the conditioning stimulus, and also between evoked potentials in different regions of the CNS. It was suggested that iterated stimulation builds a net which reflects the temporal pattern of previous inputs, i.e. entrainment indicates the influence of afferents on nets; and by such mechanisms a multiple representational system is developed. Since wave-shapes *per se* do not propagate in the manner of impulses, it was suggested that the wave-shapes arising in neuronal nets influence the spatiotemporal patterns of spikes emitted therein, so that the pattern of postsynaptic excitation in nets which receive such spikes reflects the time course of excitation in the initial nets. Thus the evoked potential represents the temporal input to neuronal nets. It may or may not reflect the time course of events related to the outflux from these same nets. It is conjectured that neurons within a net extract the average pattern of influxes by spatial and temporal integration. John et al. in fact suggest that the wave-shape itself may be precisely what constitutes functional information for the CNS.

B.3.4. Functional Aspects of Nets

However we have so far considered only temporal aspects of the firing patterns of neurons related to macropotentials. In recent years much data has been generated in studies of the spatial aspects of neuronal interaction, particularly in the sensory projection systems. Macroelectrodes have of course been used extensively in such studies but of course the fine details have been averaged out so that essentially topographical information has been acquired. In this way the gross organization of receptive fields and the representational systems within the CNS have been elucidated. What emerged from early studies was the concept of a rather static nervous system – a central core wherein all the important processing was carried out, with a well-defined localized representational system in its sensorimotor cortex, connected to its receptors and effectors by transmission lines or cables. The new concept of organization which has emerged from work with microelectrodes is much more dynamic in that even peripheral nets are seen to process signals in a complex fashion by selecting, filtering, abstracting, computing, under constant control from more central nets. Indeed hierarchies of nets execute a progressively selective response to stimuli as the map from periphery to centre unfolds.

In the visual system for example, a very detailed hierarchy has been discovered. Early works by H. K. Hartline, [53], [64], and later by Granit, [54] on the retinas and optic nerves of vertebrates had demonstrated the existence of units which responded phasically only to the onset or offset of a stimulus or to both. Such units have become known as "on", "off", and "on-off" units respectively. Figure C.43 shows some idealized spike-patterns generated by such units. In addition to these temporal aspects of peripheral discharges Hartline, [65], measured the receptive-fields of units in the Frog retina. Such fields are relatively large, up to 1 mm diameter at the retinal periphery, smaller towards the centre. Now in a Frog's retina there are about $3 \cdot 10^4$ nerve fibres, and about $1 \cdot 10^6$ receptor cells. Hartline's results suggest that some kind of spatial convergence must occur, and that a process of spatial summation must occur in retinal units. H. B. Barlow (1953) was able to demonstrate such summation. In addition he found that there was decreased sensitivity of "on-off" units towards the edges of their receptive fields. Figure C.44 shows the results of one of Barlow's experiments, in which a central and a peripheral spot are used. It will be seen that illumination of both the central and side spots results in a much smaller response than when the central spot is alone illuminated. Barlow attributed this to a process of "surround inhibition", that is the excitatory centre of the receptive field of the "on-off" unit is surrounded by an inhibitory region. This region appeared to interact linearly with the excitatory centre, suggesting spatial summation of both excitatory and inhibitory potentials. It was also noted that "on-off" units respond to a change in the spatial pattern of light, even if there is no change in the total incident light flux. Barlow concluded that there were overlapping, receptive fields in the retina, each comprising about $2 \cdot 10^3$ receptors and overlapping with about eighty other fields; that "on-off" units were more than just "off" units with an inhibitory mechanism; and on epistemological and teleological grounds that "on-off" units act precisely to detect the Frog's natural prey – Flies – so that the Frog retina as a whole functions like a computer, rejecting unwanted information, and transmitting useful information. The epistemological approach was followed by J. Y. Lettvin , H. R. Maturana, W. S. McCulloch, and W. Pitts, [95], who greatly extended the earlier findings by studying the function of retinal cells in response to light and dark patterns moved in the visual field against different backgrounds. The retinal cells studied were the "ganglion" cells. It was shown that they comprised five classes: edge detectors, convex edge detectors, changing contrast detectors, dimming detectors, and darkness detectors. The first class evidently corresponds to Hartline's "on" units, the third class to the "on-off" units, and the fourth class to the "off" units. However the second and fifth classes, the convexity and darkness detectors have no earlier counterparts. Lettvin et al. were able to exhibit a structural correspondence between the different classes and different kinds of ganglion cells, especially concerning dendritic organization. They found that such cells formed five populations uniformly distributed in a retinal layer, with many overlapping receptive-fields. The axons of cells in each class were found to end

in separate layers within nets of the central portion of the Frog's nervous system, the "rectum". There was a topological map from retina to tectum. Lettvin et al. concluded that the main function of the Frog retina is to analyse stimulus patterns in terms of qualities, and that the intrinsic nature of the retinal units as well as their patterns of interaction is vital for such a task.
Detailed information concerning the nature of interaction mechanisms is provided by the work of Hartline and F. Ratliff, [66], on the eye of the Horseshoe Crab-*Limulus*. It was found that activity in any one optic nerve fibre could be elicited only by illumination of the receptor unit giving rise to the fibre – i.e. unlike the Frog retina there is no spatial summation in the *Limulus* eye. However, the discharge of Spikes by any given receptor unit is inhibited by illuminating neighbouring receptor units. Such an interaction was termed lateral inhibition. Moreover under steady illumination the inhibition exerted by one unit on another is a linear function of the rate at which the unit emits spikes, (see Figure C.45). The intensity of illumination inhibition falls off with increasing distance between units. Such factors can easily be shown to lead to "on" and "on-off" spike bursts. Lateral inhibition appears to be present in the spinal cord where it is known as "reciprocal inhibition" ([125], [37]), in the vertebrate retinas, [90], in somatic-sensory cortex, [108], auditory cortex, [78], [85], and visual cortex, [72], [73], and [10].
Hubel and Wiesel's results are of particular interest since they indicate that contrary to the discoveries of Lettvin, Maturana, et al. in frog retinas, interaction properties rather than unit properties *per se* are of primary importance for visual integration. The system studied was that of the cat. It comprises layers of differentially sensitive receptors, the "rods" and "cones" innervating a layer of "bipolar" neurons which in turn synapse with ganglion cells. From here signals proceed up to the optic nerve to a diencephalic relay net, the Lateral Geniculate nucleus, and thence to the visual or "striate" cortex, located in that region of the occipital pole known as area 17 (see [124], [118]). Because of the small size of the "bipolar" neurons, it has not yet been possible to record spikes from them. Electron micrograms, [113], show however that between the rods and cones and "bipolar" neurons there is a regular feltwork of interacting processes – bipolar dendrites, "horizontal cells" and so forth – making a complicated plexus within which some kind of image sharpening and intensifying probably occurs. The "bipolar" neurons interact with ganglion cell dendrites and "amacrine" cells in another plexus which also receive centrifugal fibres from cephalic nets. Figure C.46 shows in schematic form some of these details. The ganglion cells are somewhat larger than the bipolars and the spike patterns they emit were recorded by Hubel and Wiesel. Two classes of ganglion cells were found, those with "on" centres and surround inhibition, and those with "off" centres and surround inhibition. None of the more complicated classes found by Lettvin, Maturana et al. were found in the cat's retina. In the lateral geniculate nucleus two similar classes were found usually with pronounced inhibitory surroundings and such that a specific optic nerve fibre excited a maximum response from the cell under observation. A precise topographical representation of retinal fields was

found in the geniculate body.
In area 17 no cells were found with concentric "on" or "off"-centred receptive-fields, but a variety of new classes. In fact the cells seemed to fall into two broad general categories which Hubel and Wiesel called "simple" and "complex". Simple cells have the property that their receptive fields can be partitioned into excitatory and inhibitory regions, separated by straight lines. Stimuli such as slits, bars, and edges are likely to be the most patent stimuli for such cells. Moreover the stimuli must be of fixed position and orientation in the receptive-field. Complex cells in general cannot be partitioned simply into excitatory and inhibitory regions and they respond to slits, bars, and edges or constant orientation located anywhere within the receptive-field. Figure C.47 shows typical responses from a simple and a complex cell. In neighbouring regions of cortex, in areas 18 and 19, sometimes known as "secondary visual areas", Hubel and Wiesel found yet another category of cells which they called "hypercomplex". These are cells which respond to moving slits, edges, and bars, and to more complicated patterns containing corners, niches, and the like, all with constant orientation. However the terms "excitatory" and "inhibitory" if used to characterize the receptive-fields of hypercomplex cells must be understood in a more abstract way than hitherto, since the effects of regions do not summate spatially in any simple sense. Figure C.48 shows the response of such a hypercomplex cell. In addition to these differing cell classes, Hubel and Wiesel found a very important general property of all simple and complex cells, and of almost all hypercomplex cells, namely that cells which are close neighbours almost always have the same receptive-field orientation. Furthermore regions or "columns" of constant orientation extend from surface to white matter with walls normal to the cortical layers. Within any one column in area 17 there are both simple and complex cells, within any column in areas 18 and 19 there are hypercomplex cells of differing orders of complexity. Finally, a given small area of the retina is represented many times in area 17, in columns representing many orientations.
These results are of considerable importance in that they provide us with some insight into the nature of the "local" mechanisms of functional organization of neuronal nets. There is evidently great convergence onto bipolar neurons and ganglion cells from receptors, leading to overlapping receptive-fields of considerable extent. Although there is no direct evidence of lateral inhibition anywhere within the vertebrate visual system, it is plausible that horizontal and amacrine cells in the retina are responsible for the appearance of surround inhibition in ganglion receptive-fields. There appears to be an approximately 1:1 map from the ganglionic layer to the lateral geniculate, however the map from hereon to area 17 is probably many:many since the simplest assumption needed to explain the responses of simple cortical cells is that they act as coincidence detectors of inputs from the lateral geniculate. Similarly complex cortical cells can be considered to be coincidence detectors of inputs from simple cells, and hypercomplex cells as detectors of various combinations of complex cell activity. Sholl's data (op.cit.) suggests that cells which are interconnected tend to be nearest neighbours, and

since there appears to be a prominent vertical component in dendritic orientation, [99], there is an anatomical substratum for the columnar organization. Indeed the column emerges as the functional unit of the cortex. To quote Hubel and Wiesel:

"Far from being a mere aggregation of cells with common characteristics, the column emerges as a dynamic unit of function. In the visual areas the columnar system can probably be looked on as a Solution to the problem of dealing with three independent variables – two to specify position on the retina and the third for receptive field-orientation – in a structure, the cortex that is in a sense two-dimensional."

It is significant that a columnar structure was also discovered in somatic-sensory cortex, [108]. Indeed columnar organization may be a feature of many cortical nets. Hubel and Wiesel's discoveries provide some indication of how stimulus coding takes place in the vertebrate visual system. It is noteworthy that the coding seems to result from analysis of local attributes of stimuli. Nothing definite is said concerning how these attributes are integrated into a holistic representation of the stimulus. Indeed nothing is said concerning how codes for such local attributes modify the maintained activity found in neuronal nets ([91] and [22]) which is held to be fundamental to the integration of nets. One possible reason for this is that Hubel and Wiesel have worked entirely with animals under barbiturate anaesthesia. As a result many of the observed cells which responded to retinal stimulation were not spontaneously active. A. D. J. Robertson (1965) found that even very low doses of barbiturate reduced the number of spontaneously active cells observed, and generally increased the number of cells which responded to retinal stimulation. Moreover receptive-field organization was greatly altered and stabilized by the anaesthetic. J. C. Lilly, [97], recorded with macroelectrodes the spatial distribution of neocortical activity in anaesthetized animals and found that the effect of anaesthesia was to isolate functional regions from one another in the cortex, to simplify their activity, and to prevent interactions between neighbouring regions. Burns et al. working with anaesthetized preparations could evoke responses only from spontaneously active cells. Most cells gave "on-off" responses to visual stimuli, superimposed on a background of maintained activity. It was suggested that signals were presented to the visual system as short-lasting deviations from the mean firing rate of cortical cells. Kuffler et al. ([91]) reached similar conclusions concerning the retina.

B.4. Discussion

As we have seen, macroscopic studies of neuronal anatomy and physiology lead to the view of the nervous system as a complex of ordered nets arranged in a three level hierarchy, with many circuits within and between the constellations of neurons we have called supernets. On the average such nets are specific; there are for example well-defined laminar and columnar structures in cortical nets. EEGs, ECoGs, and D.C. potential records indicate that there is an underlying dynamical activity generated within supernets. Feedback inhibition undoubtedly plays a very important role in maintaining the activity which is controlled both by external and internal signals, especially from nets of the lower levels – the reticular systems. The EEG serves as a measure of the general level of the dynamic activity of supernets, the D.C. potentials as measures of the stationary states of such activity, i.e. as long-time averages, and the ECoG and evoked responses as measures of the incoming and outgoing fluxes in nets. There are two broad categories of evoked responses. One is associated with tonic aspects of the nervous system, is related to the gross spontaneous activity, and is probably generated in the same corticothalamic circuits. The other is the modality-specific response associated with phasic aspects of the system, generated in different circuits.
The view of the nervous system which results from microscopic studies of single cells and their interactions, fills in the details necessarily obscured by macroscopic studies. However there are also significant differences, the resolution of which is a difficult problem. Studies of isolated cells have established the manner in which cells respond to stimuli and emit spikes. In general there is a monotonic relation between the rate of spike-emission and the currents generated in somato-dendritic regions of the neuronal membrane. The role of transmitter substances, of EPSPs and IPSPs, and of ions in determining such responses is becoming known. There are definite relationships between spike-patterns and macropotentials, at least on the average. It is probable that such potentials are the algebraic sums of EPSPs and IPSPs seen through some kind of low-pass filter, and that D.C. potentials reflect the states of balance between EPSPs and IPSPs that constitutes the stationary states of the dynamical system. Studies of the interactions between cells and of the phasic responses of cells to patterned stimuli have established that there is a specific organization of columnar nets, maintained by nearest-neighbour interactions, spatial summation and inhibition. There are topological maps between many nets and from receptive-fields. Studies of the effects of anaesthesia indicate that tonic responses may also be present in the cells of "physiological" animals, although it is not clear that the precise columnar organization is then completely preserved.
Concerning the neuronal coding of signals, one apparently general operating mode is via pulse rate modulation, so that signals are deviations from steady-state activity. The view that the evoked response signals by modulating the stationary balance between EPSPs and IPSPs comprising the. D.C. potentials is consistent with this. However pulse rate modulation is only one possibility. Receptive field experiments

indicate that the specificity of the individual neuronal response is a very important signal. These responses occur as spike bursts, and it is possible that in physiological preparations, variations in such parameters as the frequency of bursts, number of spikes per bursts, together with cell specificity, and response latency all constitute meaningful signals. One view is that almost all changes in the spatio-temporal patterns of firing within neuronal nets convey information. *Per contra* it has been suggested that nets average over incoming fluxes by processes of spatial and temporal summation and integration within cells, plus diffusion between cells, and that the shape of the evoked response – the average pattern of influxes and outfluxes – is the important signal. Certainly receptive-field experiments indicate that signals set-up locally within nets are important for detecting and abstracting qualitative features of stimuli. Neither how these features are integrated into a holistic representation of stimuli, nor how such representations are eventually coded and stored, is yet known. The columnar organization which has been discovered points to such structures as being the "functional units" of the nervous system, and it is worth noting that columnar coordinated plus orientation presumably supplied by dendritic geometry, serve to represent a three-dimensional Euclidean manifold. We conclude this survey of neuronal nets by observing that what we have outlined and probably oversimplified is only one aspect of a much broader vista. We cannot hope to understand how neuronal nets function without knowing how they came to be, and what controls their continued existence, i.e. neurochemistry, genetics, embryology, and so forth. Nor have we looked at the functional or informational aspects of nets, for example learning, recognition, and so forth. It has been said that the great need is for a physiological theory of thought, and that a way must be found to reconcile the phenomena of perceptual generalization, the stability of memory, and the instabilities of attention, [67]. Perhaps the electrophysiology of neuronal nets will one day help to answer these questions. Certainly the problems will occupy neurophysiologists, mathematicians and others for a very long time. Meanwhile perhaps the most appropriate description of neuronal nets is still that by C. S. Sherrington, [141].

"A scheme of lines and nodal points, gathered together at one end into a great raveled knot, the brain, and at the other trailing off to a sort of stalk, the spinal cord. Imagine activity in this shown by little points of light. Of these some stationary flash rhythmically, faster or slower. Others are travelling points, streaming in serial trains at various speed. The rhythmic stationary lights lie at the nodes. The nodes are both goals whither converge, and junctions whence diverge, the lines of travelling lights...
(The brain is) an enchanted loom where millions of flashing shuttles weave a dissolving pattern, always a meaningful pattern though never an abiding one; a shifting harmony of sub—patterns."

B.5. Acknowledgement

The work embodied in this review was performed with financial support from the U.S. Department of the Navy, Office of Naval Research, Physics Branch, under Contract No. N.62558-4256, for which support the author is extremely grateful. The author would also like to express his gratitude to Sir Willis Jackson, F.R.S., on behalf of the Electrical Engineering Department, Imperial College of Science and Technology, for the warm hospitality he has enjoyed as an Academic Visitor, while this review was written. Many helpful discussions with Mr. N.W.Ellis and Mr. A.D.J. Robertson are gratefully acknowledged. Finally the author would like to thank Mrs. L.M. Talbot and Mrs. B. Broadhurst for the preparation of the manuscript and figures.

C. Figures

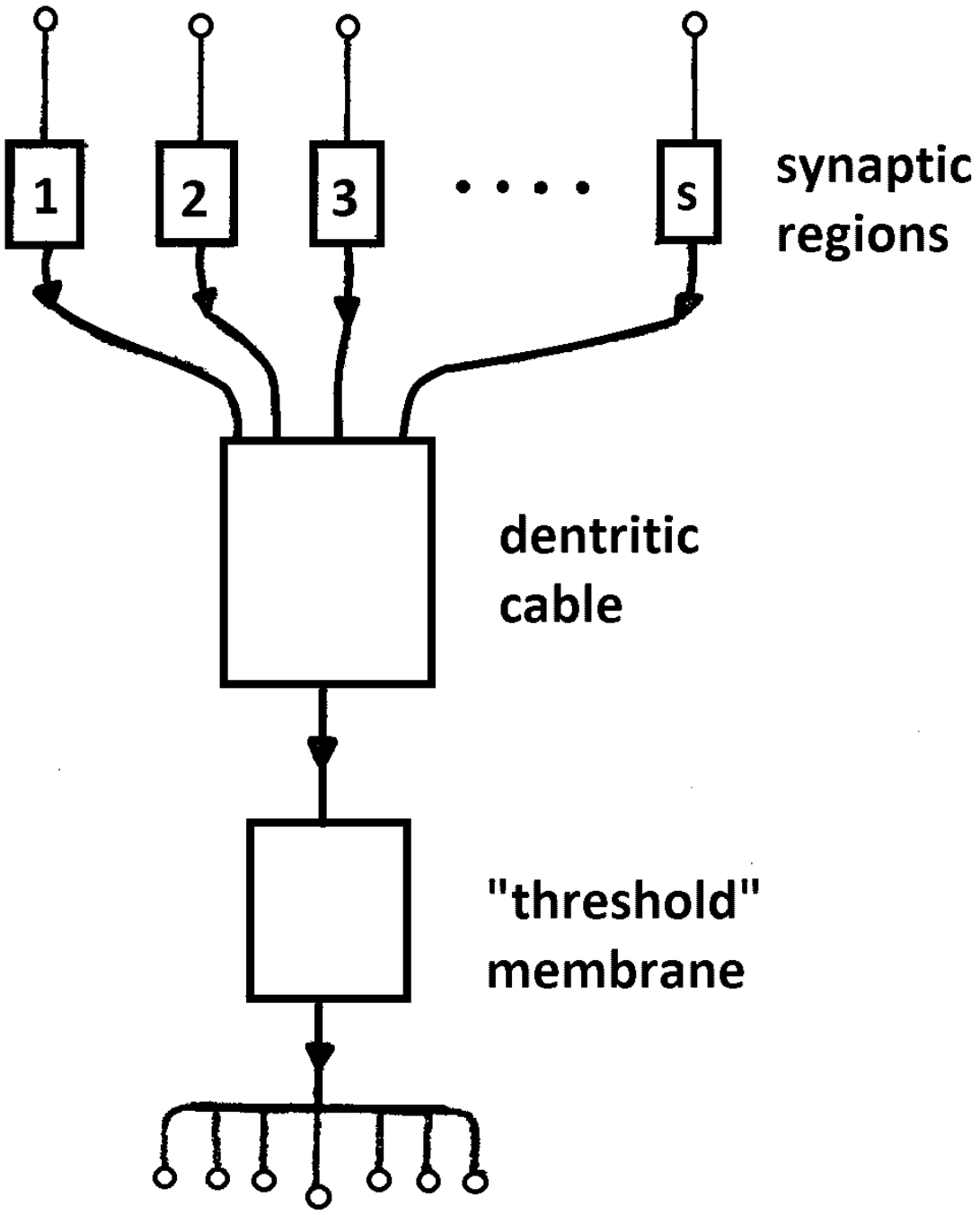
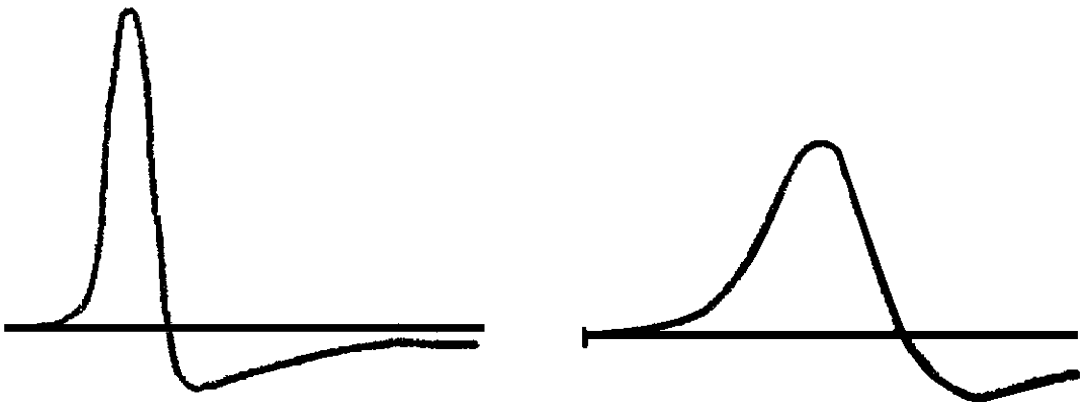


Figure C.1.: The model neuron



(a) Current impulse at sub-synaptic region

(b) "Smeared" current impulse following transmission through dendritic cable

Figure C.2.: Two current impulses – sub-synaptic region and "Smeared"

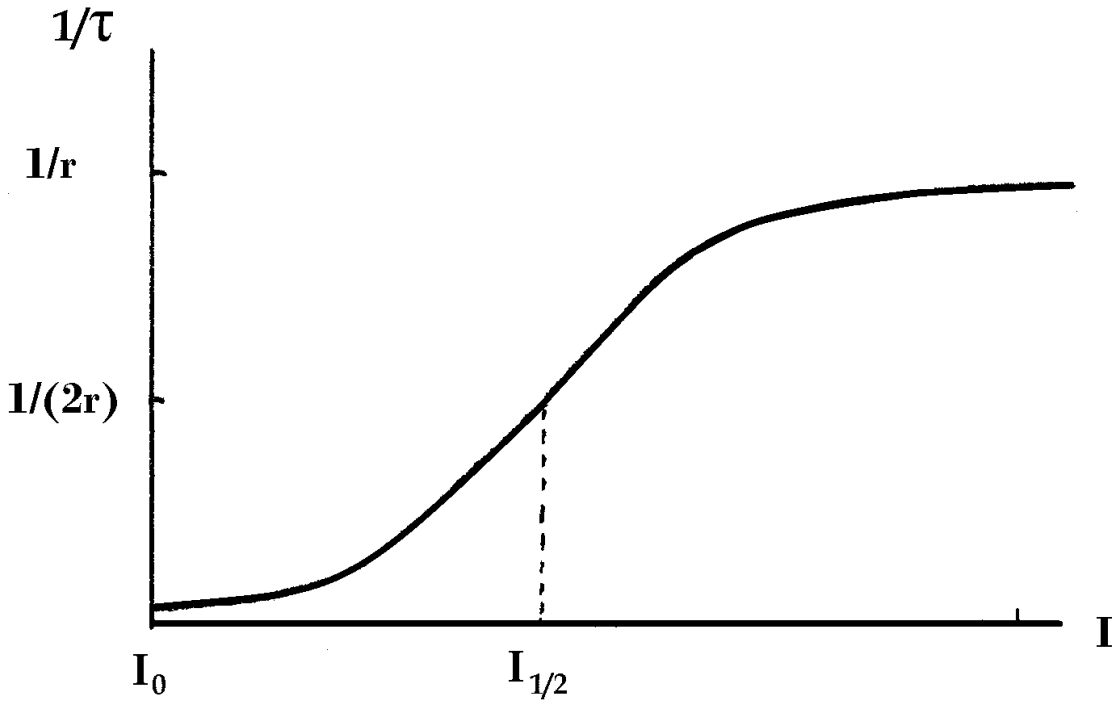


Figure C.3.: The logistic function

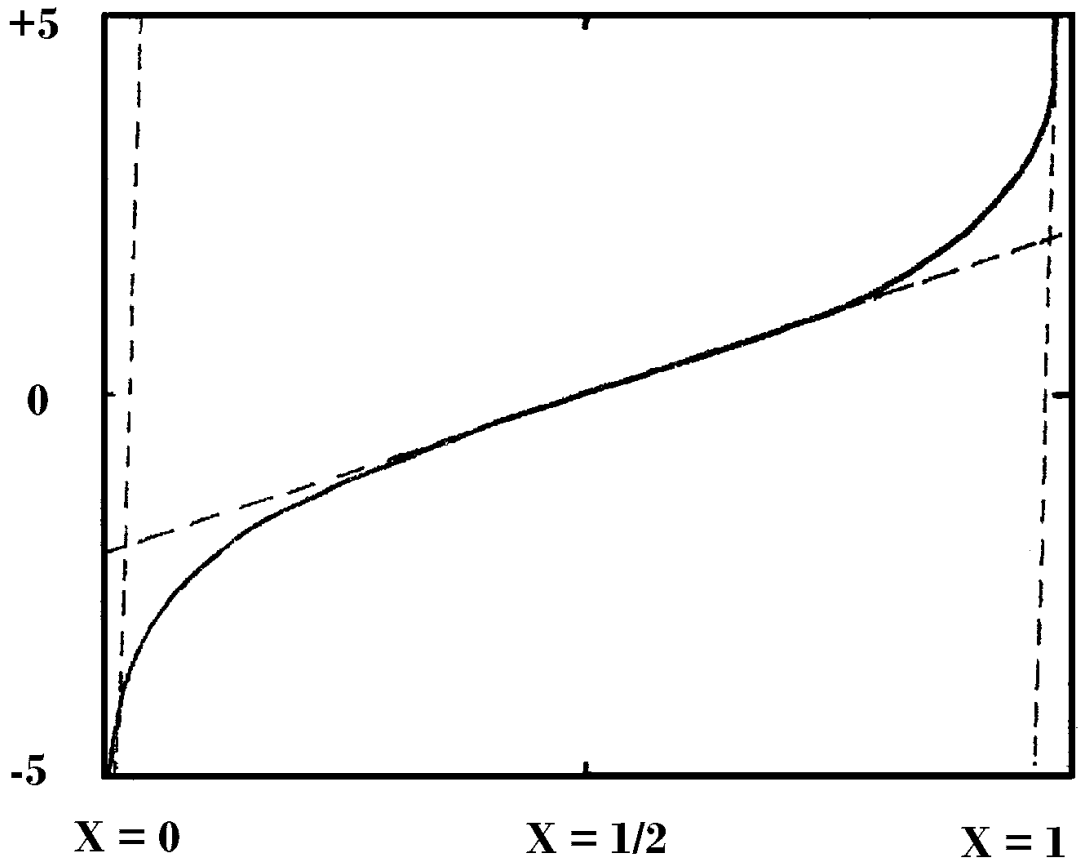


Figure C.4.: $\ln \frac{X}{1-X}$ vs. X

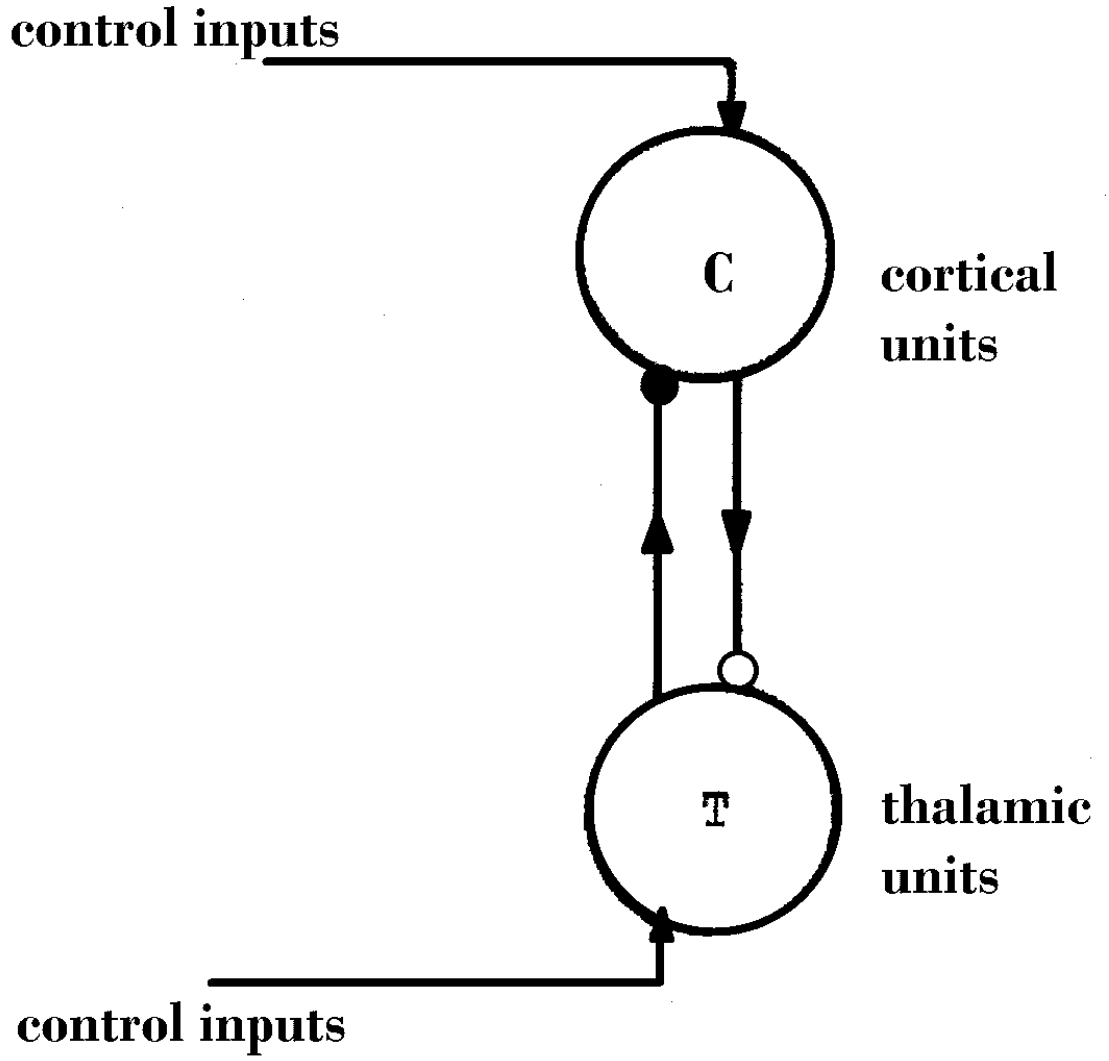


Figure C.5.: Postulated Corticothalamic circuits

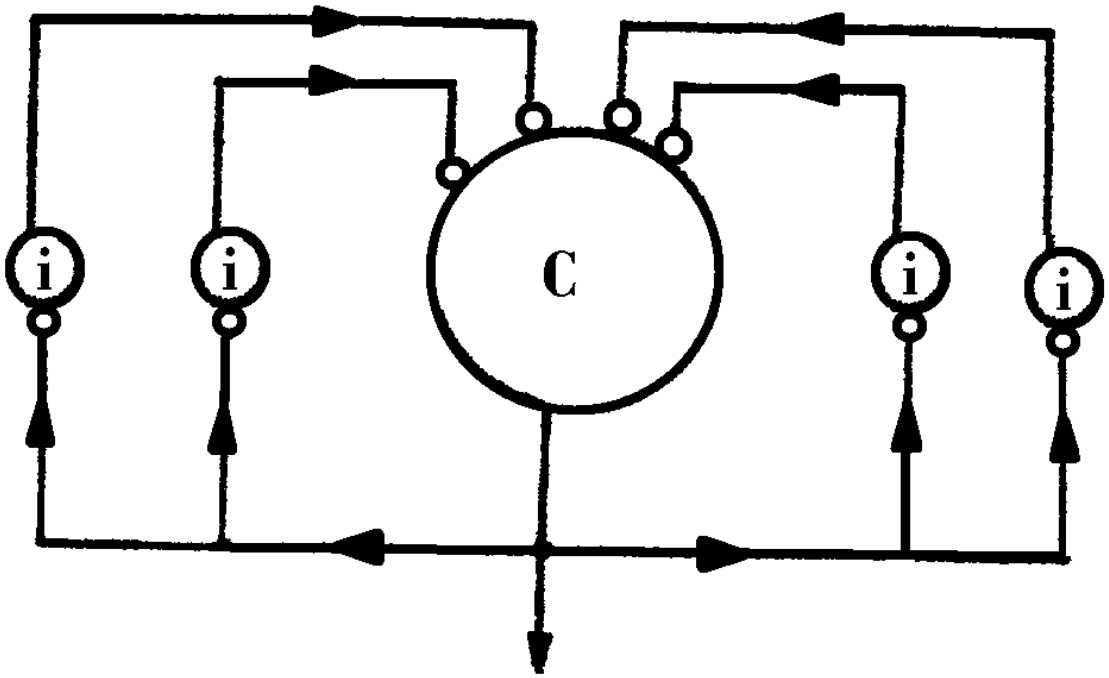


Figure C.6.: Interneuron circuits

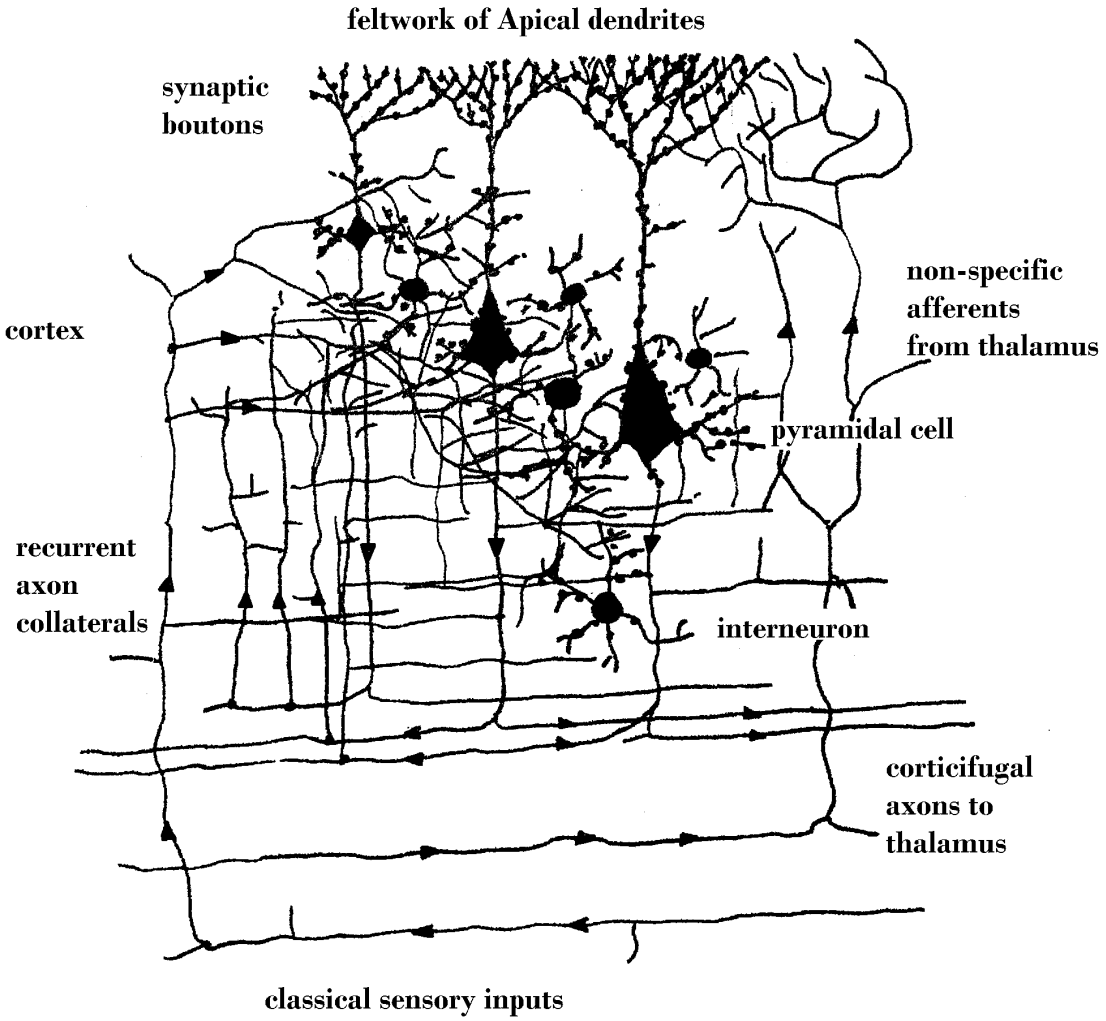


Figure C.7.: Redrawn from Cajal (1894), [123] (see Appendix B), Cortical anatomy

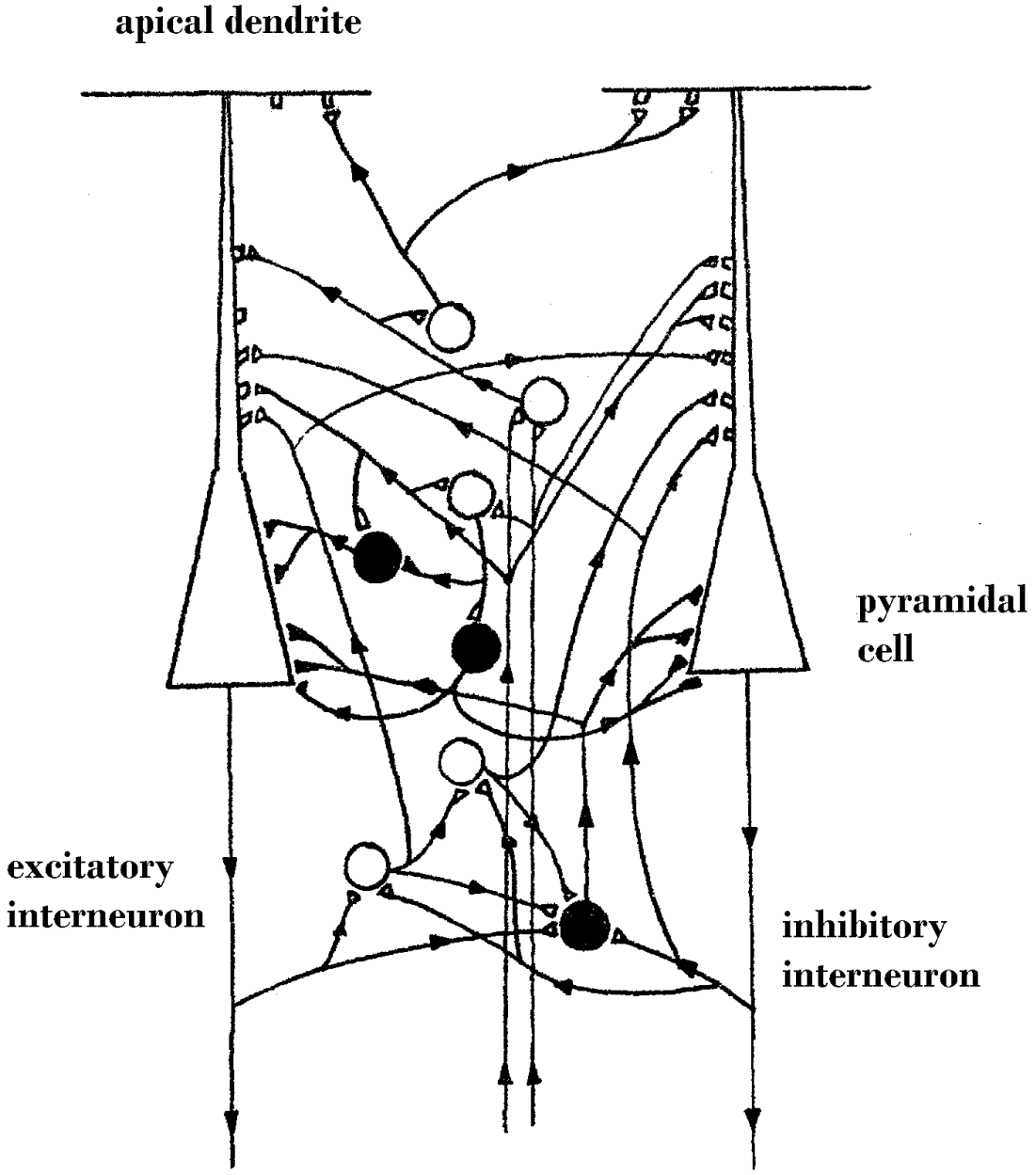


Figure C.8.: Recurrent collateral and afferent excitation and inhibition. (Redrawn from [36])

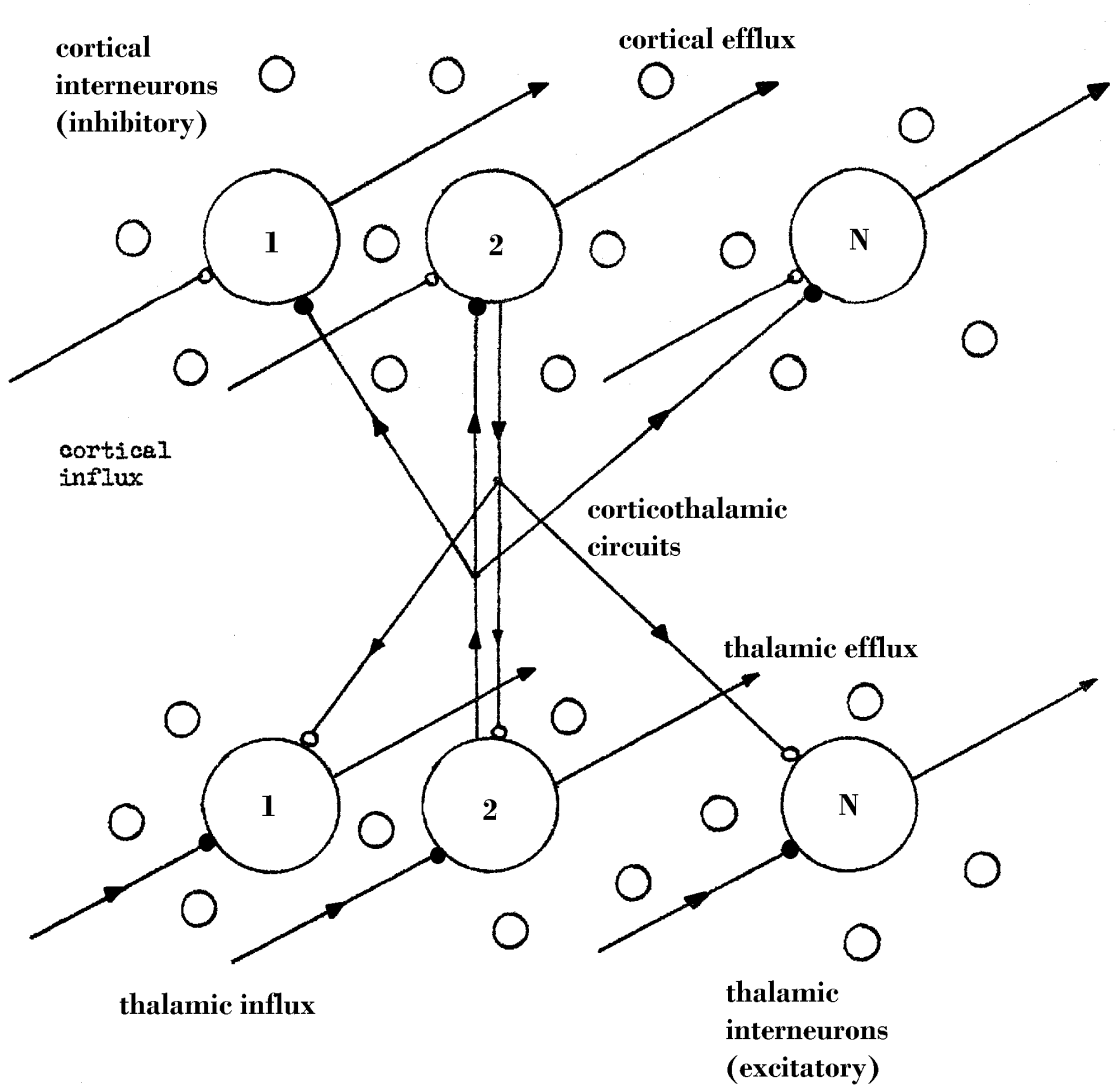


Figure C.9.: Interactions of cortical and thalamic neurons corresponding to the differential equations Equation 4.1 or Equation 4.4

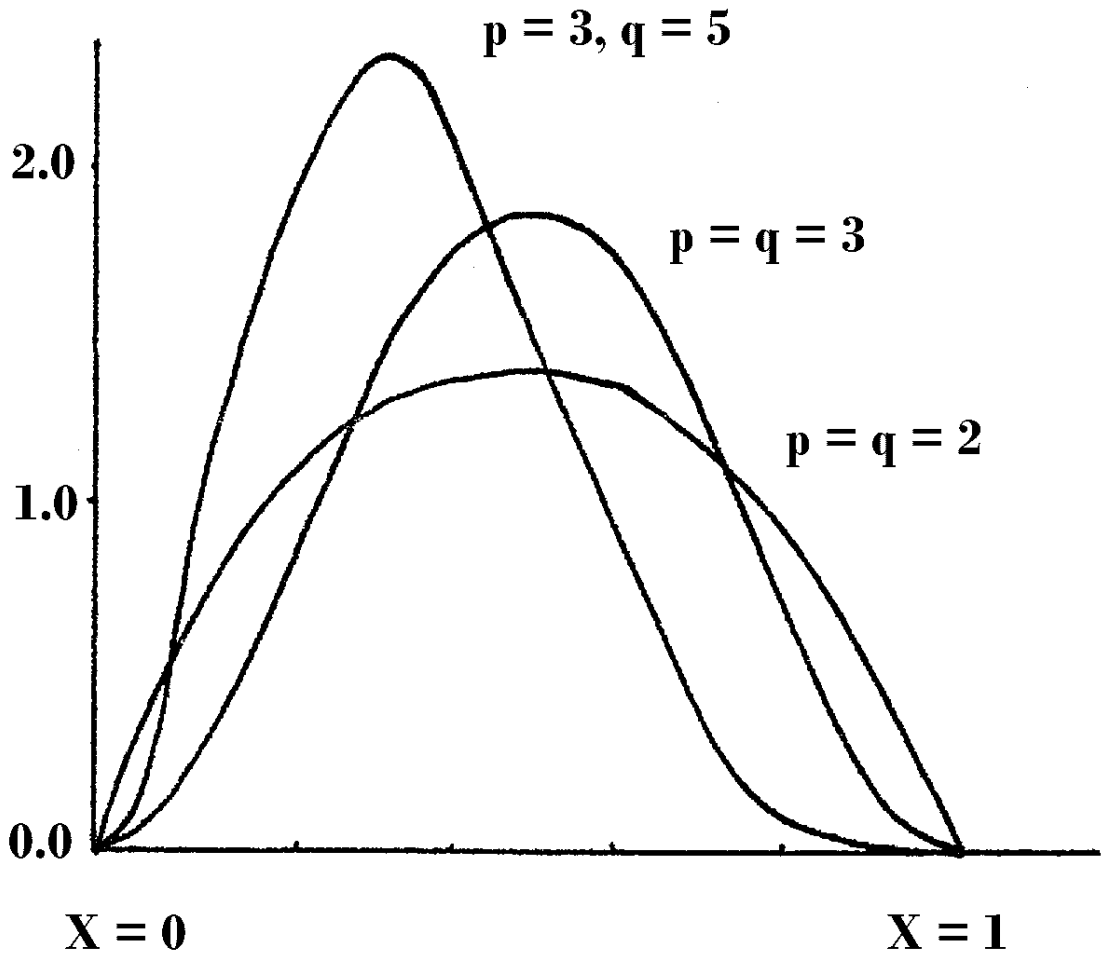


Figure C.10.: The Beta density function

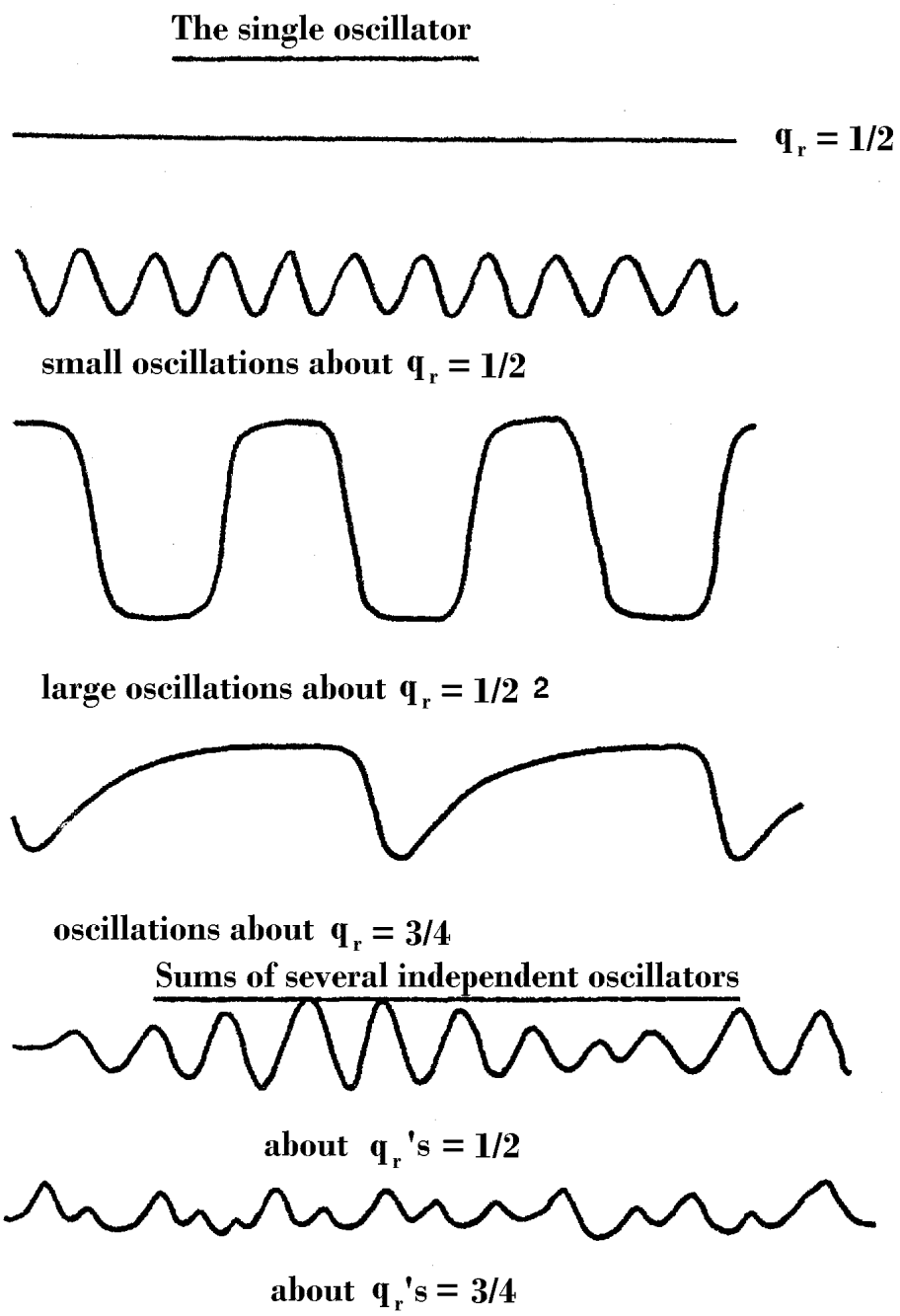
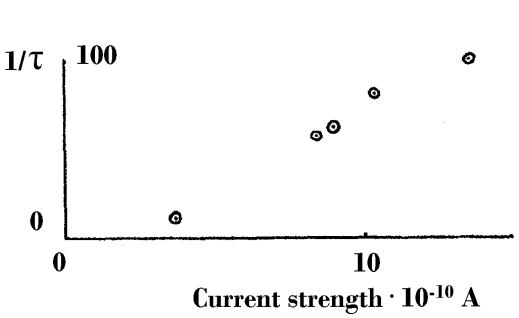
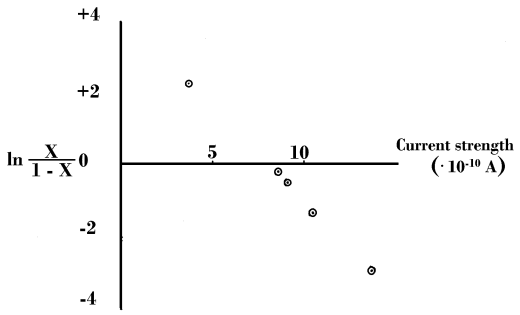

The single oscillator
$q_r = 1/2$
small oscillations about $q_r = 1/2$
large oscillations about $q_r = 1/2$ 2
oscillations about $q_r = 3/4$
Sums of several independent oscillators
about q_r's = 1/2
about q_r's = 3/4

Figure C.11.: Sensitivity oscillations



(a) Relationship of frequency, of discharge to strength of applied intracellular current (from [32])



(b) The function $\ln \frac{X}{1-X}$ vs. I for the cell in (a)

Figure C.12.: Frequency, discharge and intracellular current of a cell

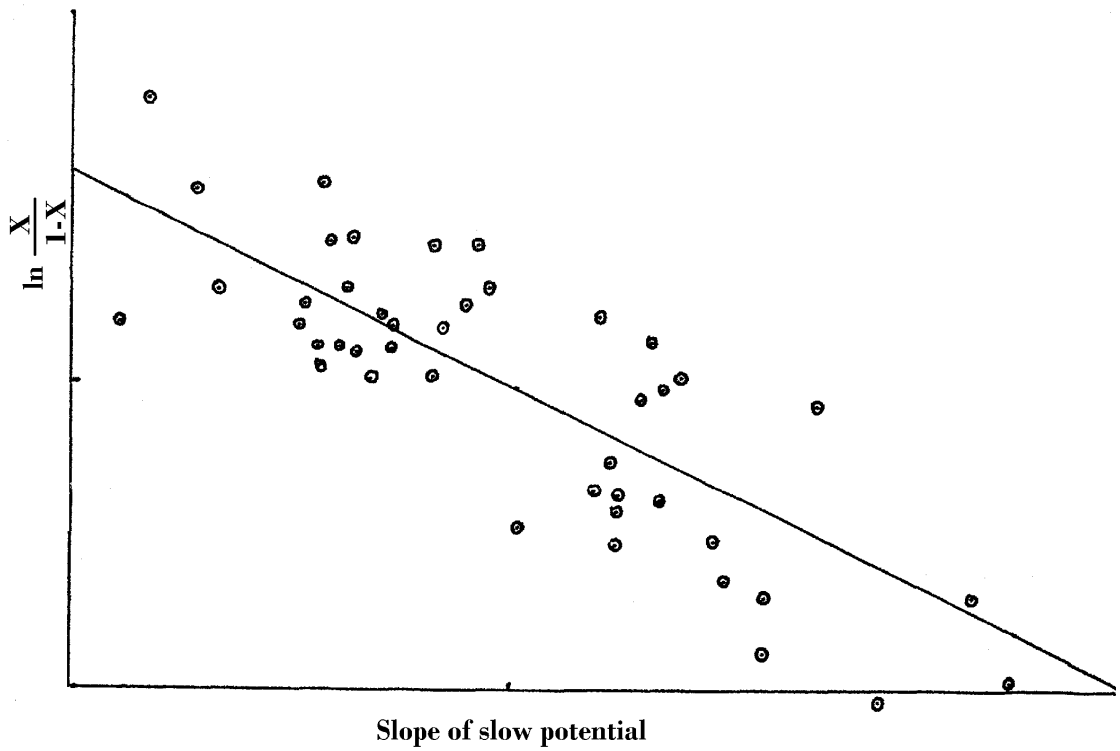


Figure C.13.: The function $\ln \frac{X}{1-X}$ vs. I estimated from extracellular experiments (see text)

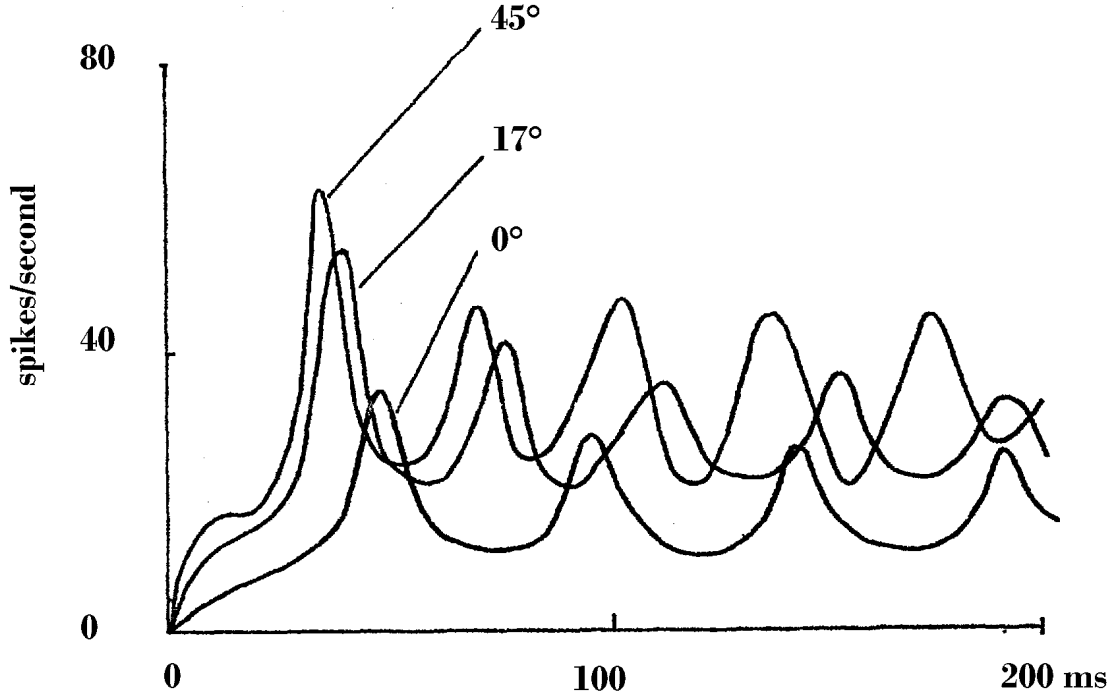


Figure C.14.: Smoothed Expectation Densities for a thalamic-somatic sensory neuron measuring joint angles (from [117])

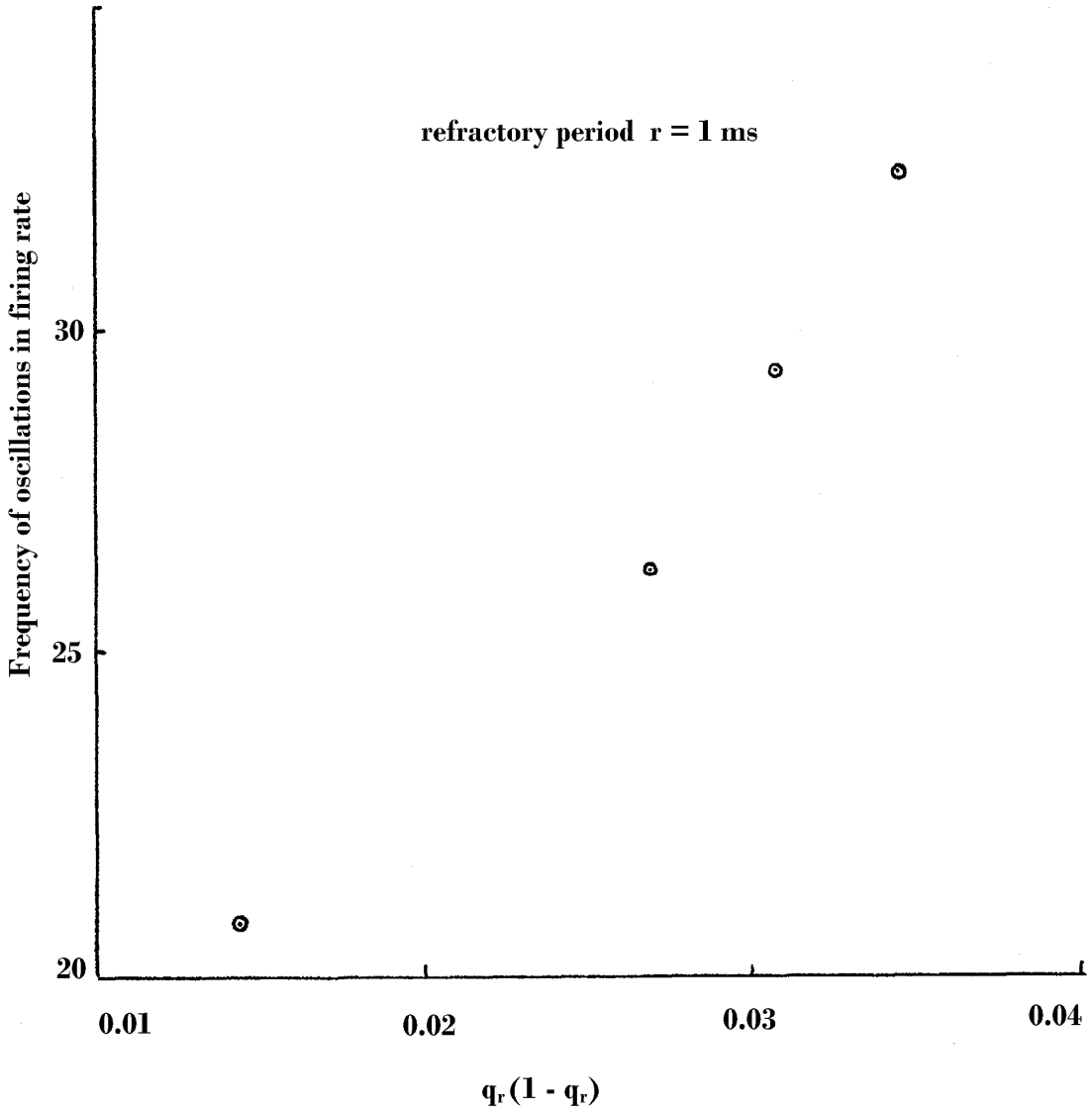


Figure C.15.: Frequency of oscillations in firing rate of thalamic-somatic sensory neurons vs. $q_r(1 - q_r)$

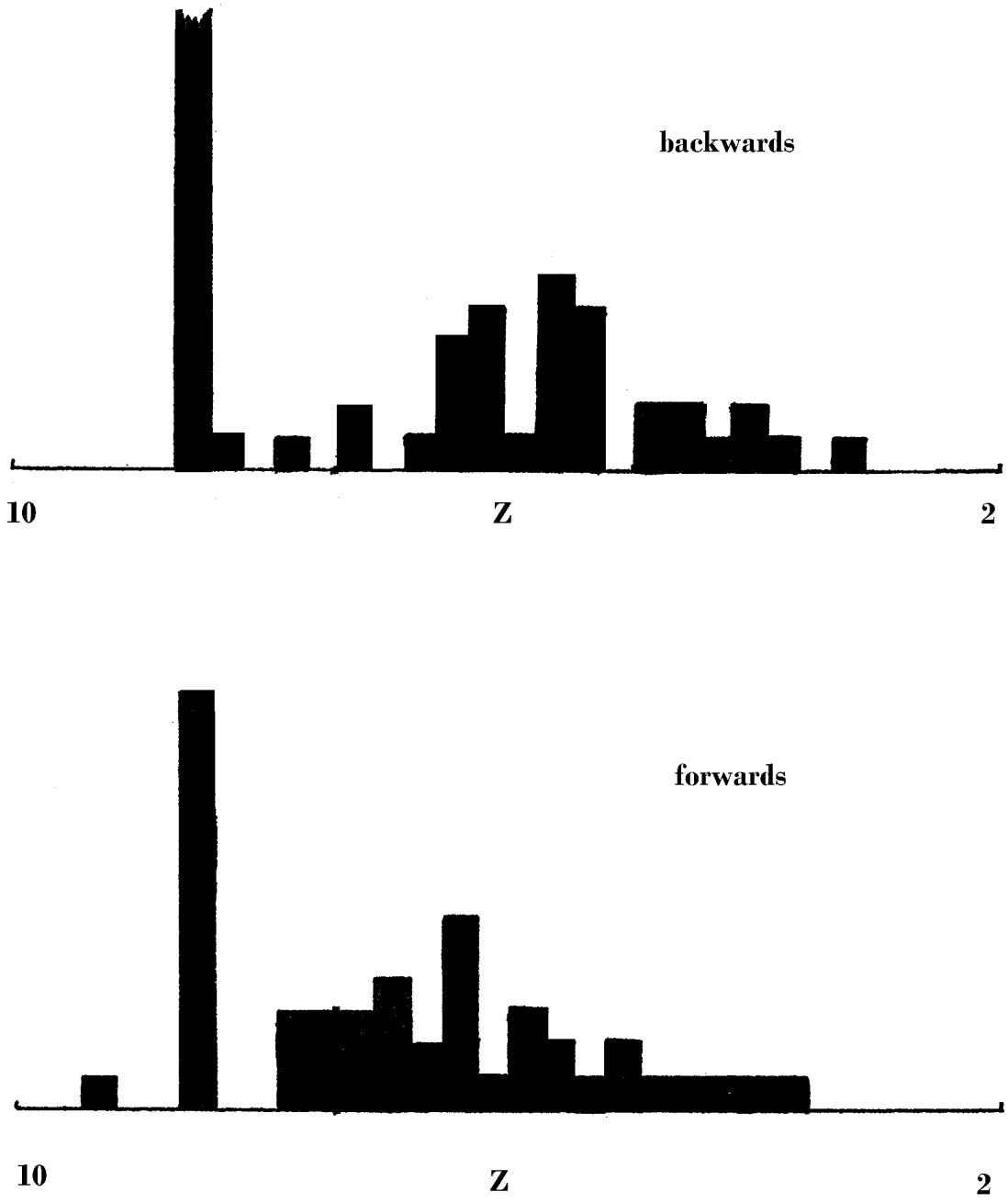


Figure C.16.: Temperature Histograms

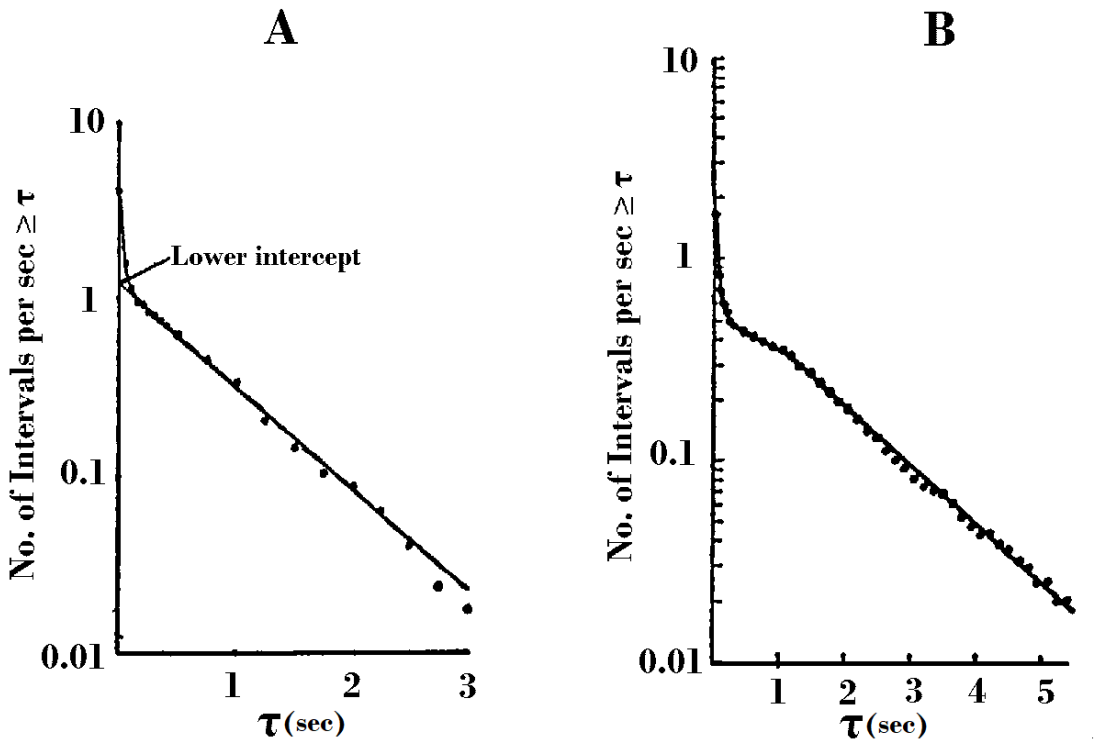


Figure C.17.: Cumulative Interval Distributions for two cortical cells (from [145])

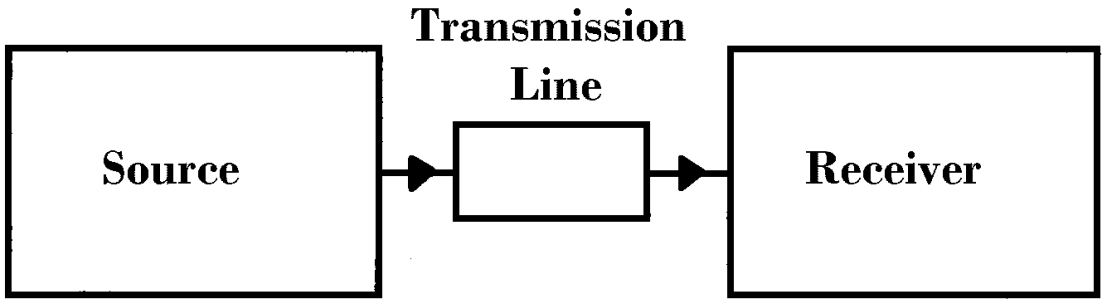


Figure C.18.: A communication system

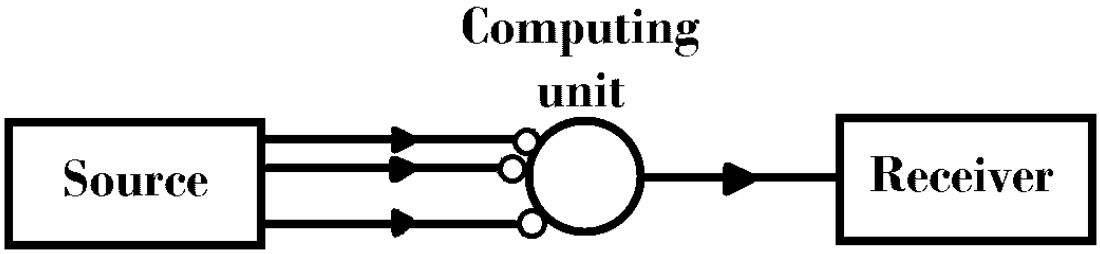


Figure C.19.: A computation system

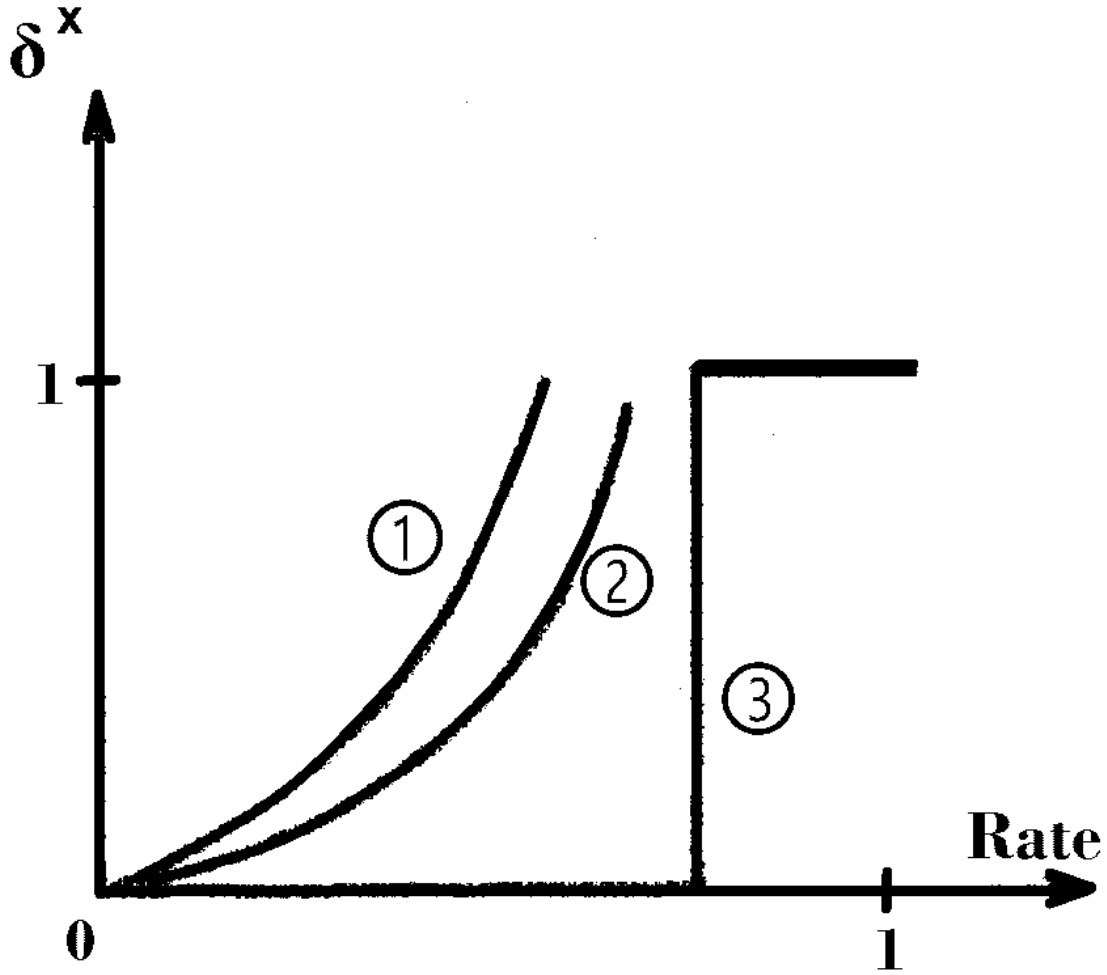


Figure C.20.: Error probability vs. Rate of transfer of information
①, ② von Neumann et al.
③ Shannon

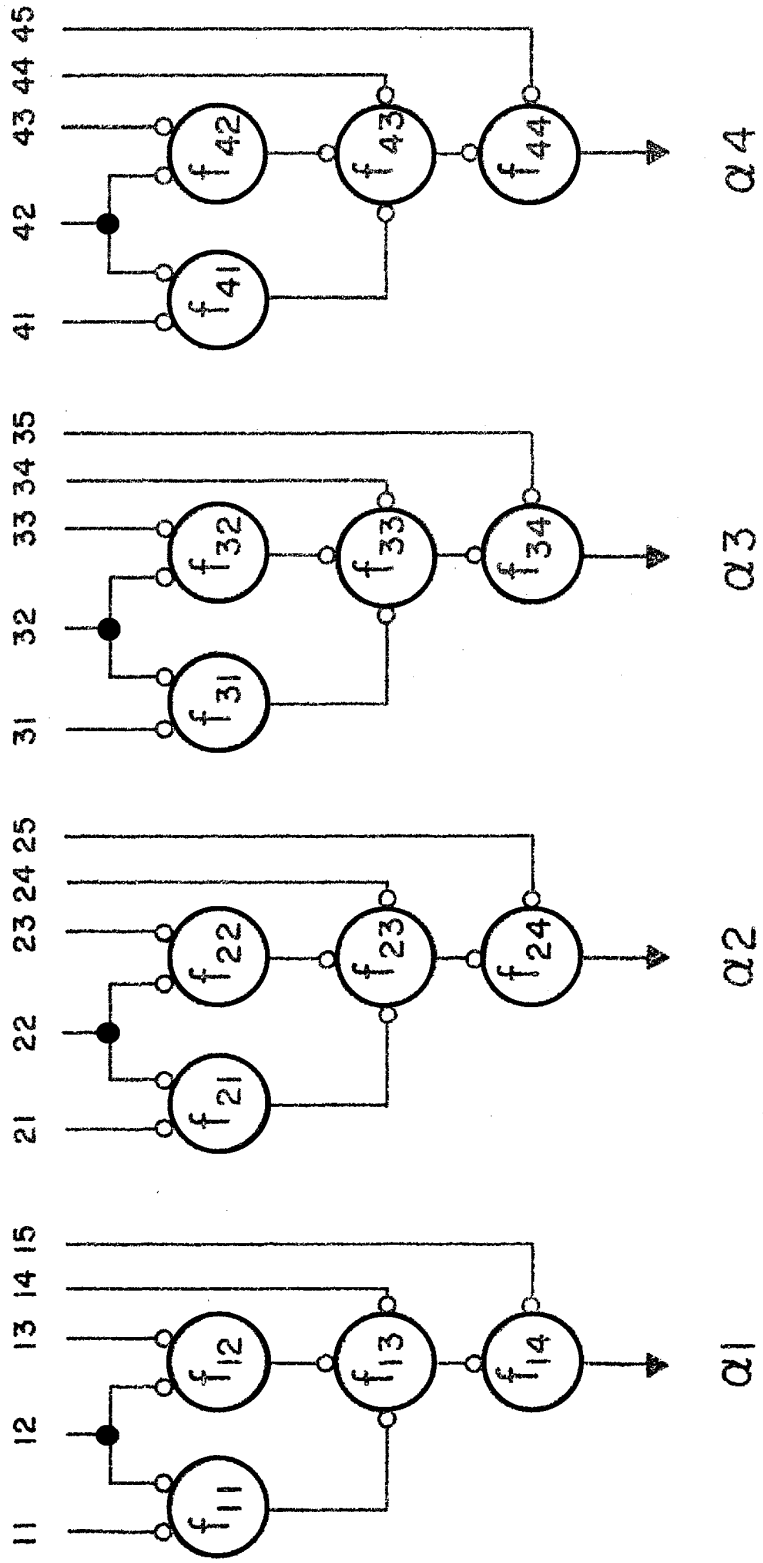


Figure C.21.: Uncoded automaton

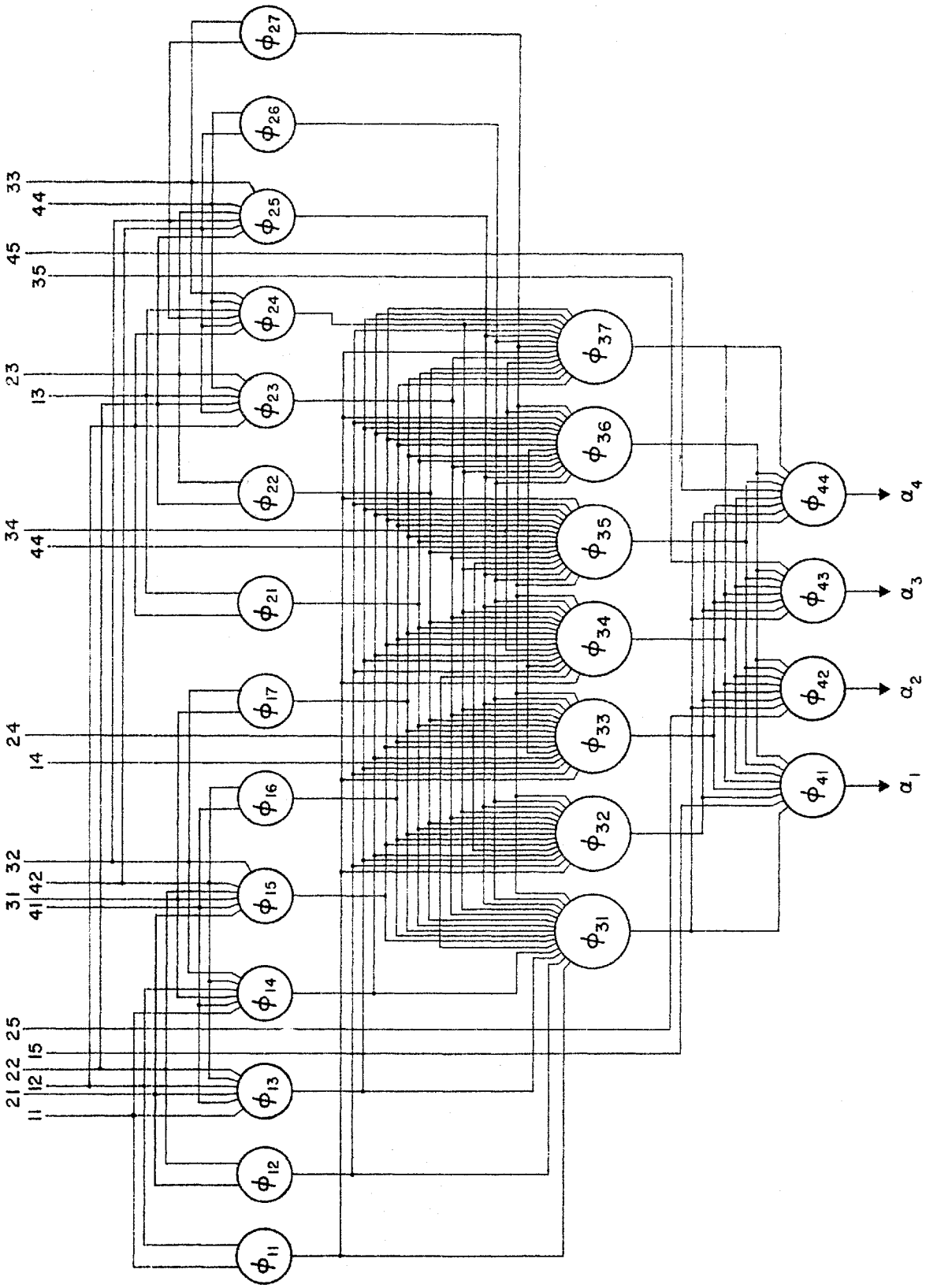
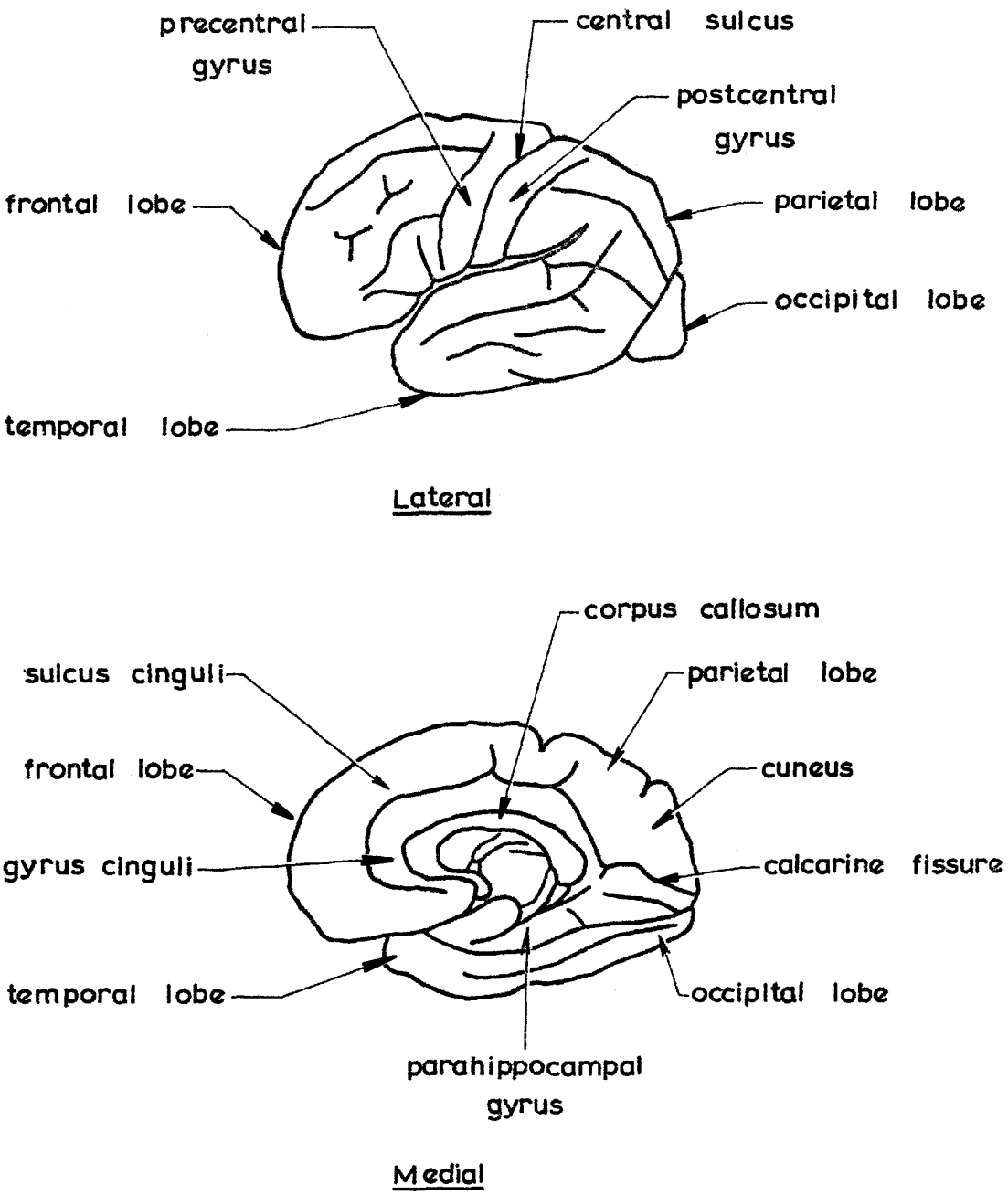


Figure C.22.: Coded automaton



(a) Fig1

Figure C.23.: Cerebral Hemispheres, from [124]

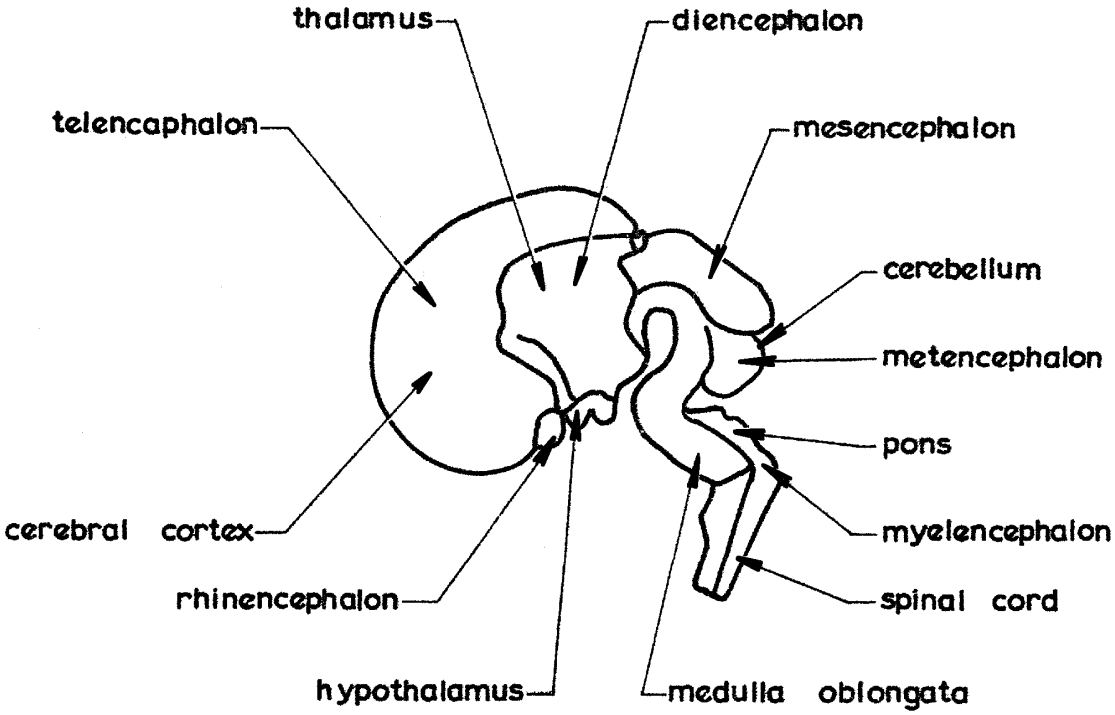


Figure 23 (Cont.).: Three months old human fetus, from [124]

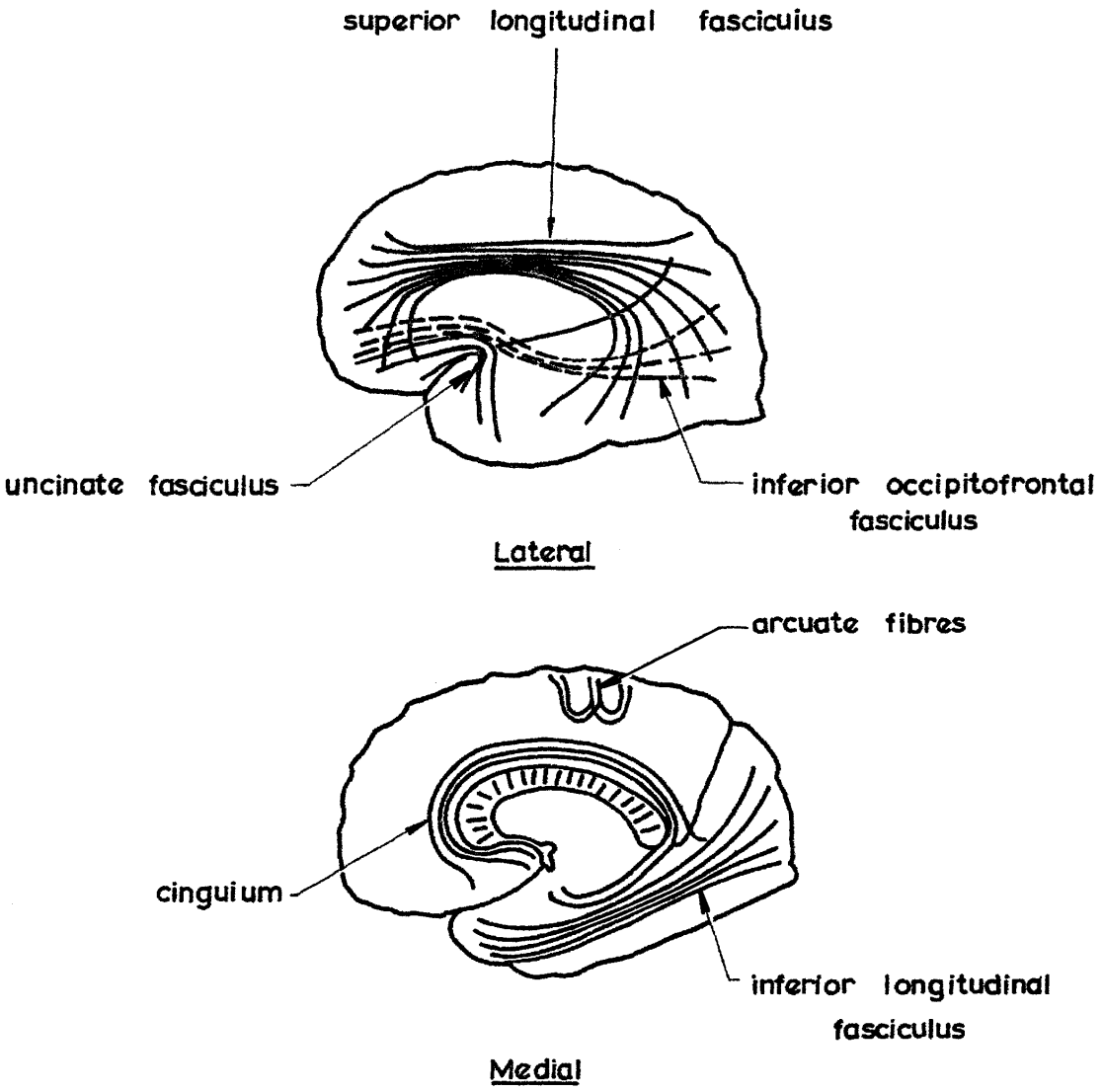


Figure C.24.: Axonal Bundles, from [124]

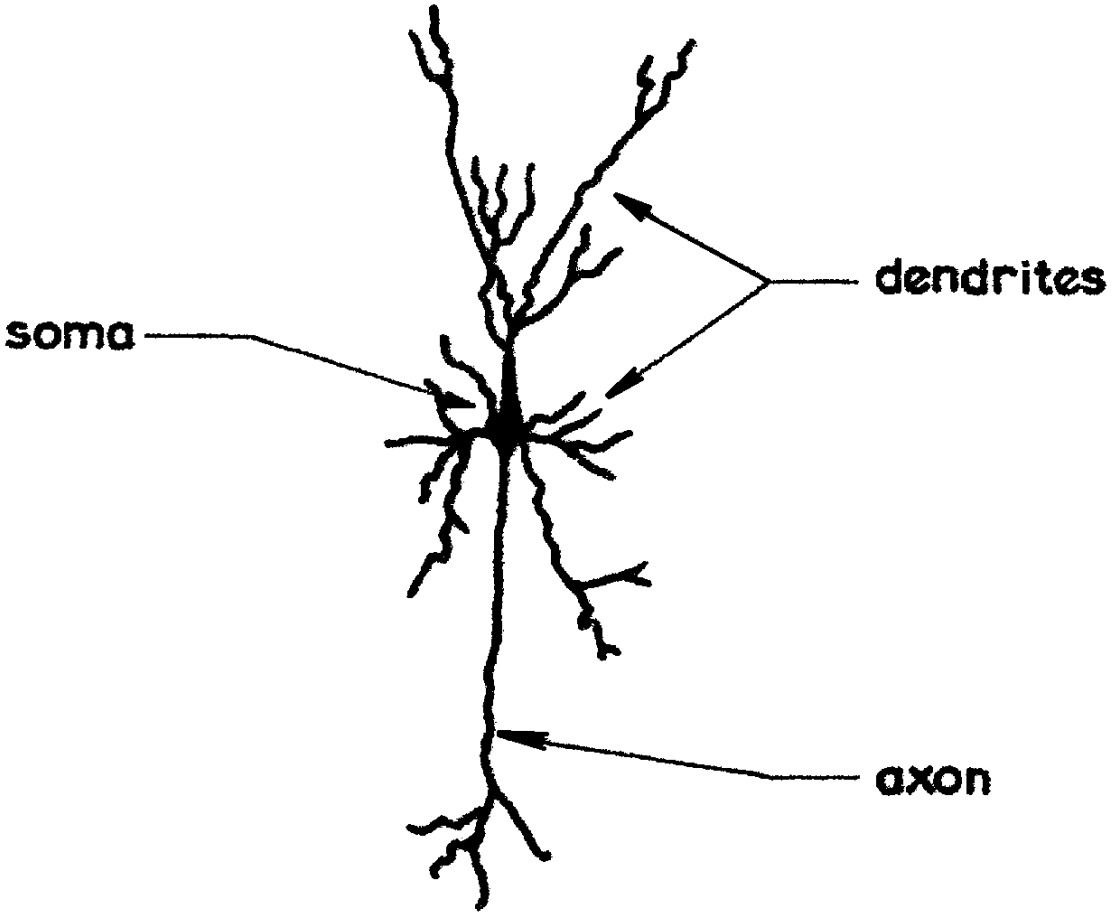


Figure C.25.: A Neuron

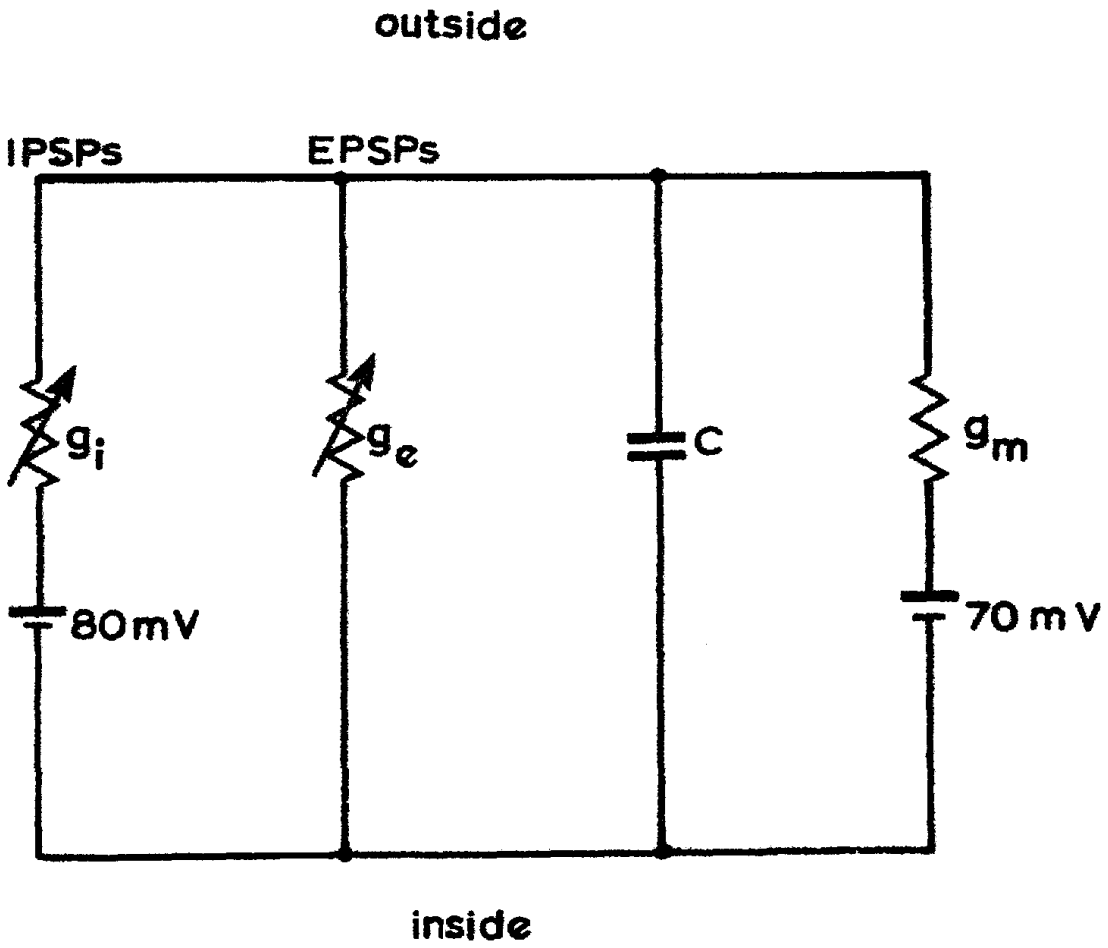


Figure C.26.: Neuronal membrane – an equivalent circuit, from [35]

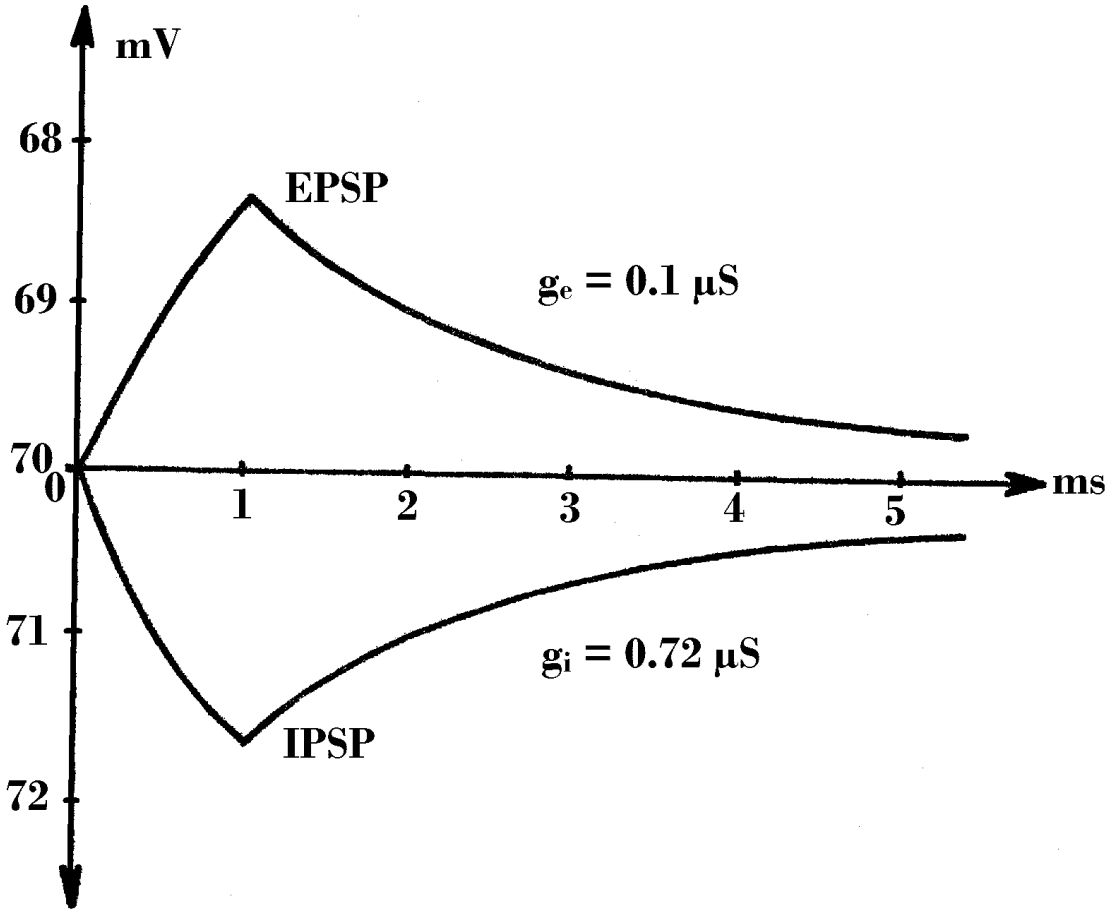


Figure C.27.: Some typical responses of the equivalent circuit

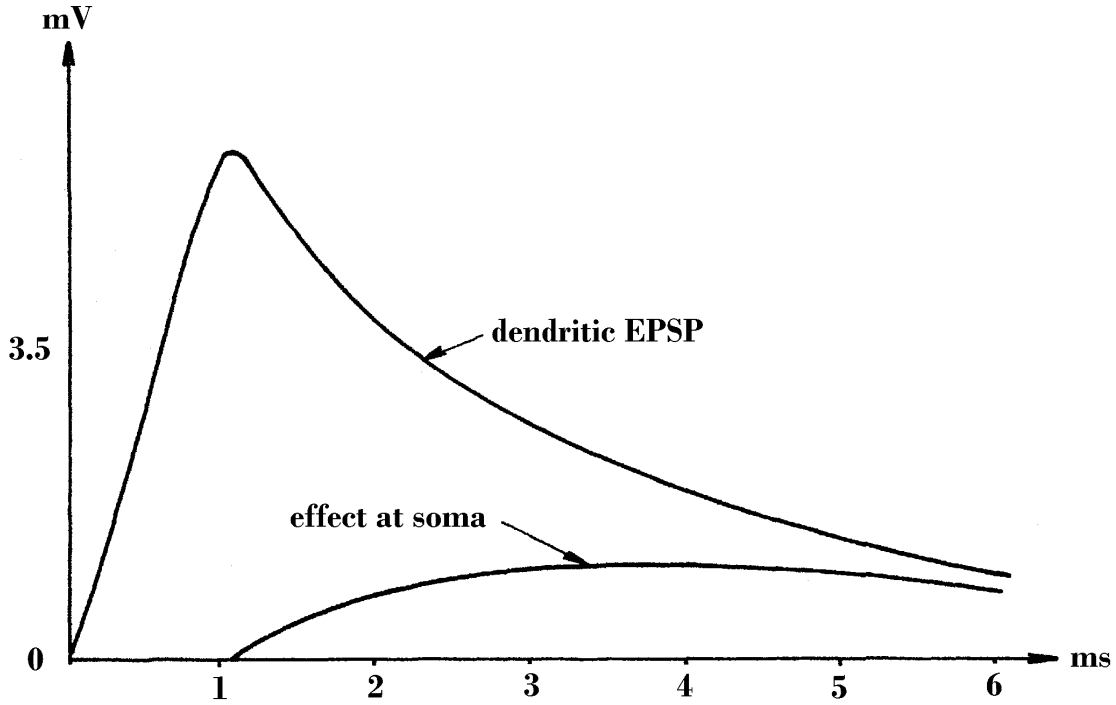


Figure C.28.: The effect of distributed parameters, from [121]

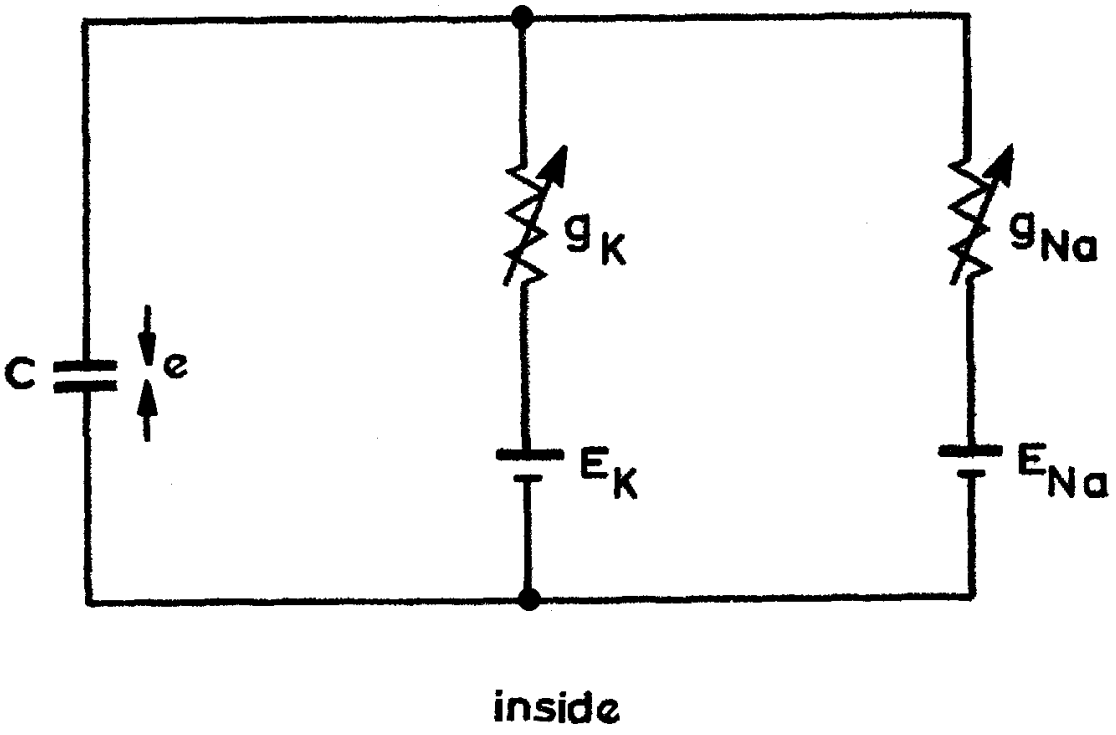


Figure C.29.: IS-membrane – an equivalent circuit, from [69]

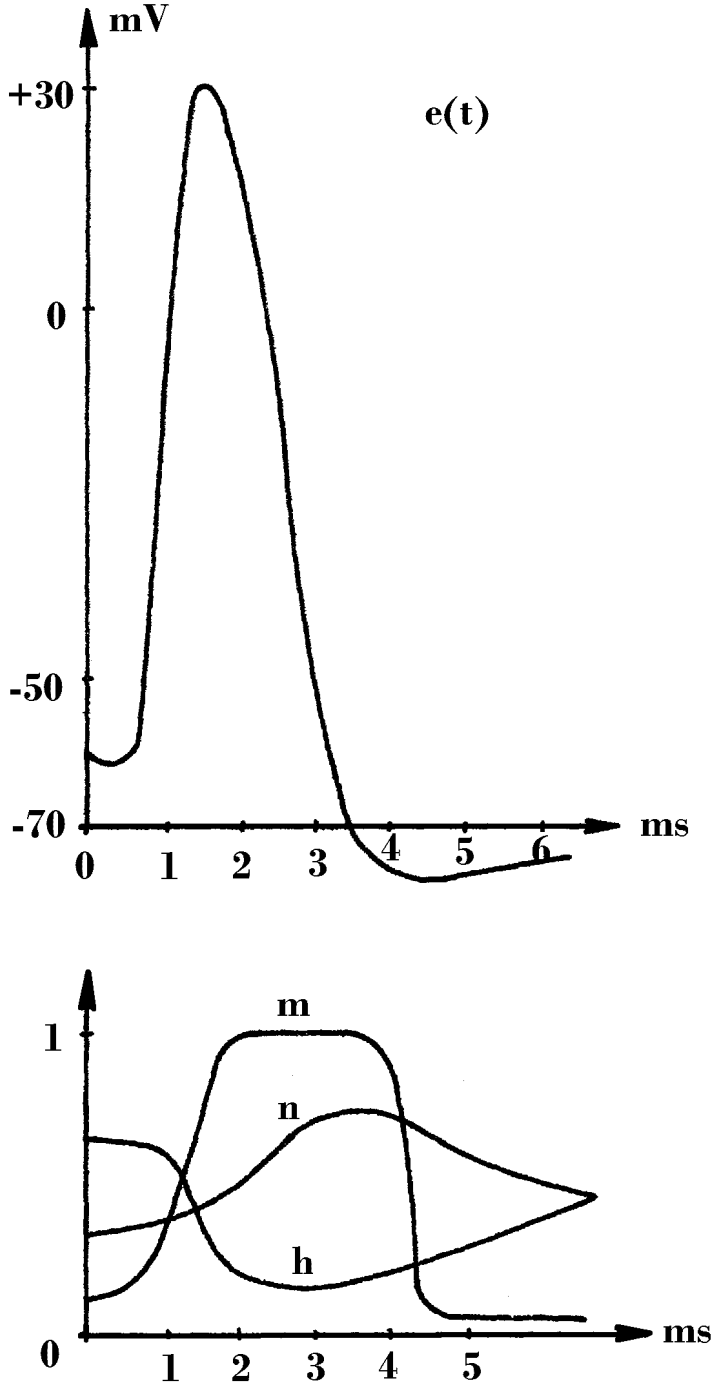


Figure C.30.: Analogue computer solutions of the Hodgkin–Huxley equations, from [44]

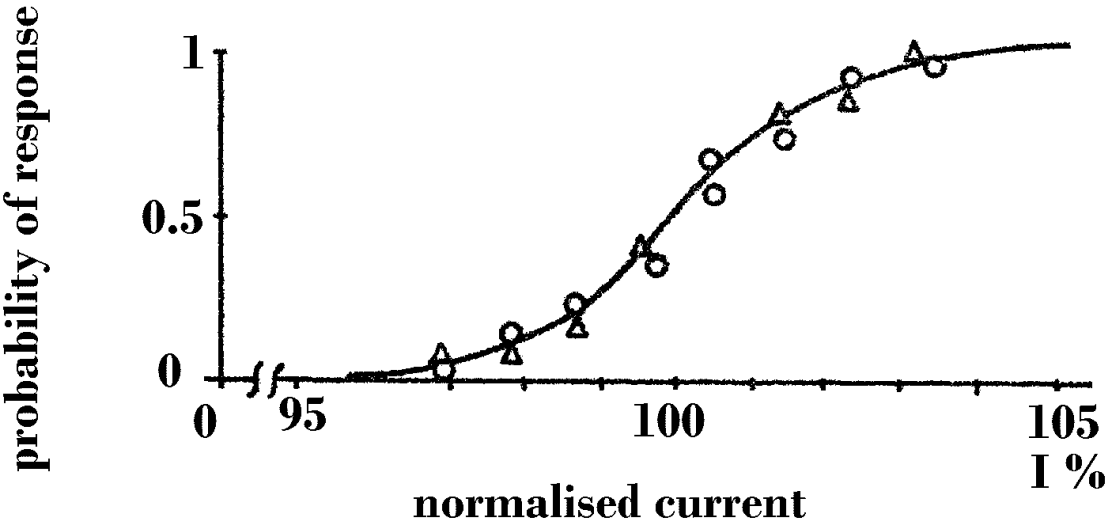


Figure C.31.: Rate-intensity curves of neuronal spike emission, from [155], [55], [68]

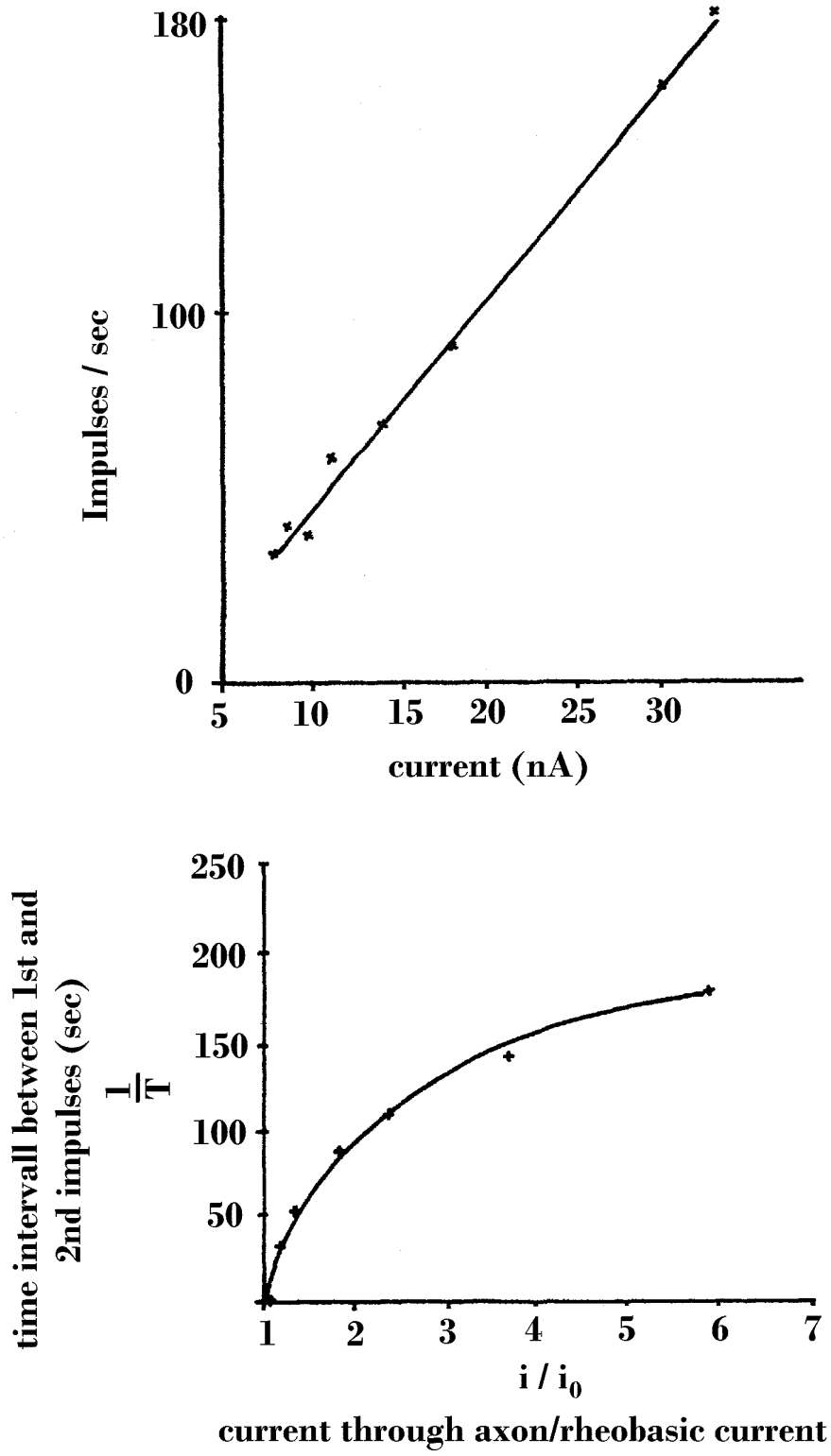


Figure 31 (Cont.).: from [155], [55], [68]

Figure C.32.: Resting EEG of a normal adult

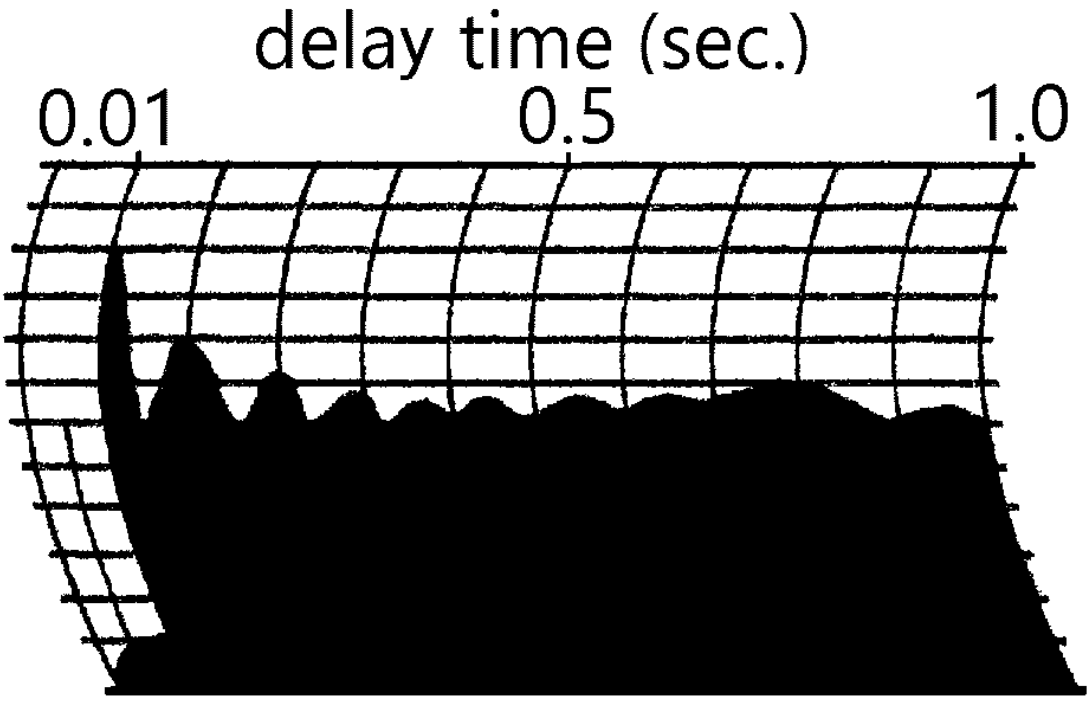


Figure C.33.: Autocorrelogram of resting EEG of normal adult, from [137]

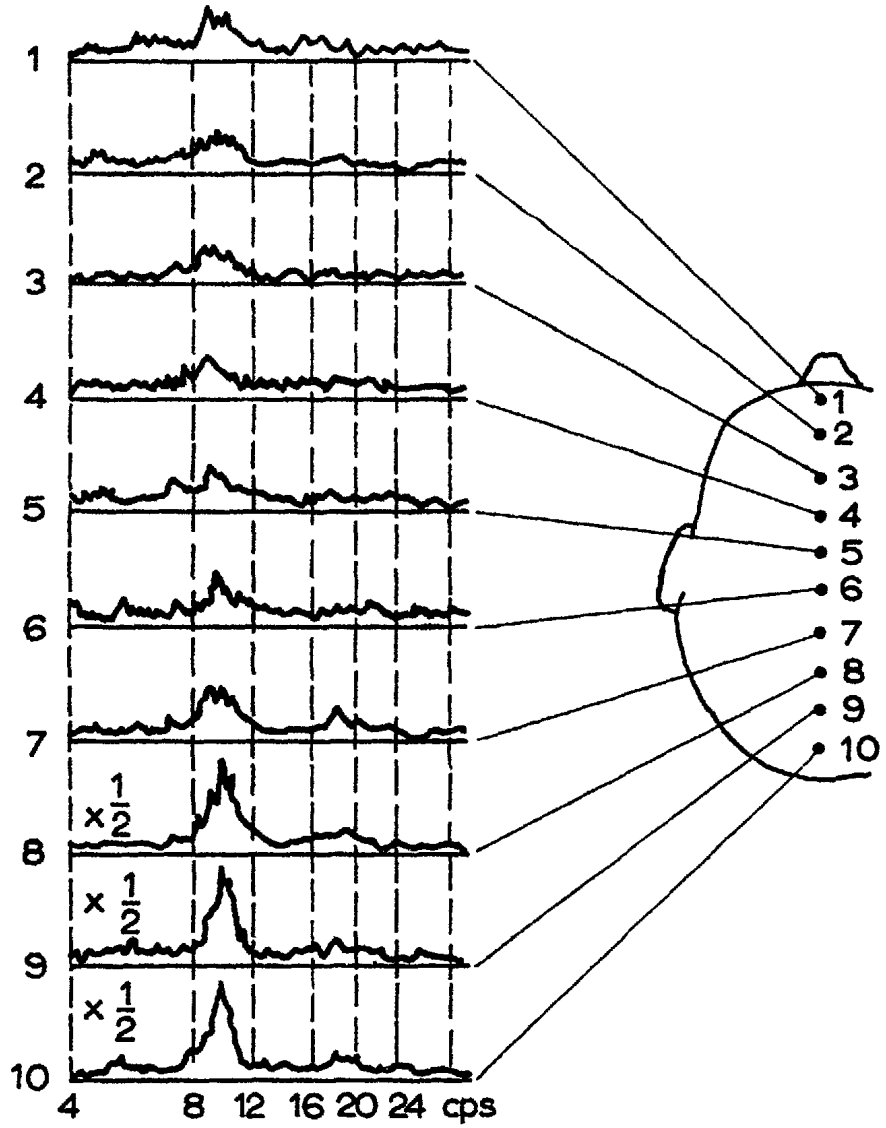
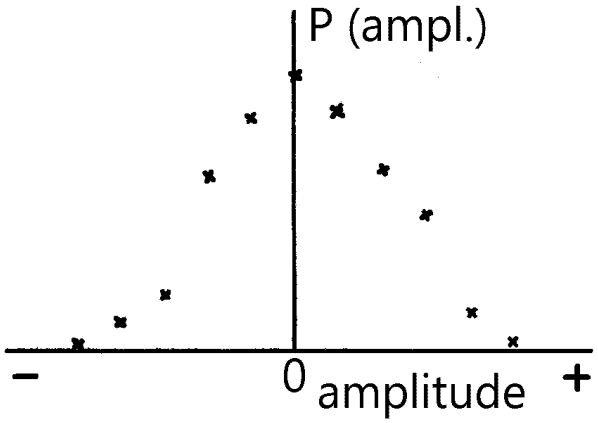
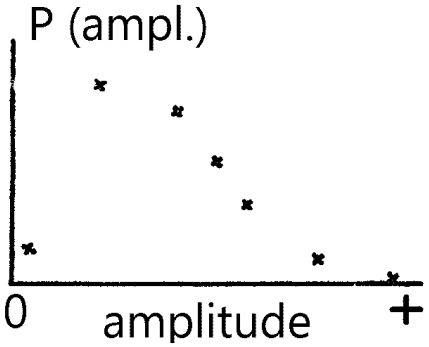


Figure 33 (Cont.).: Spectra of an adult EEG, from [148]



Probability Density of Amplitudes of
EEG of a normal adult



Probability Density of Amplitudes of EEG
Envelope of a normal adult

Figure C.34.: Probability density of amplitudes of EEG from [137]

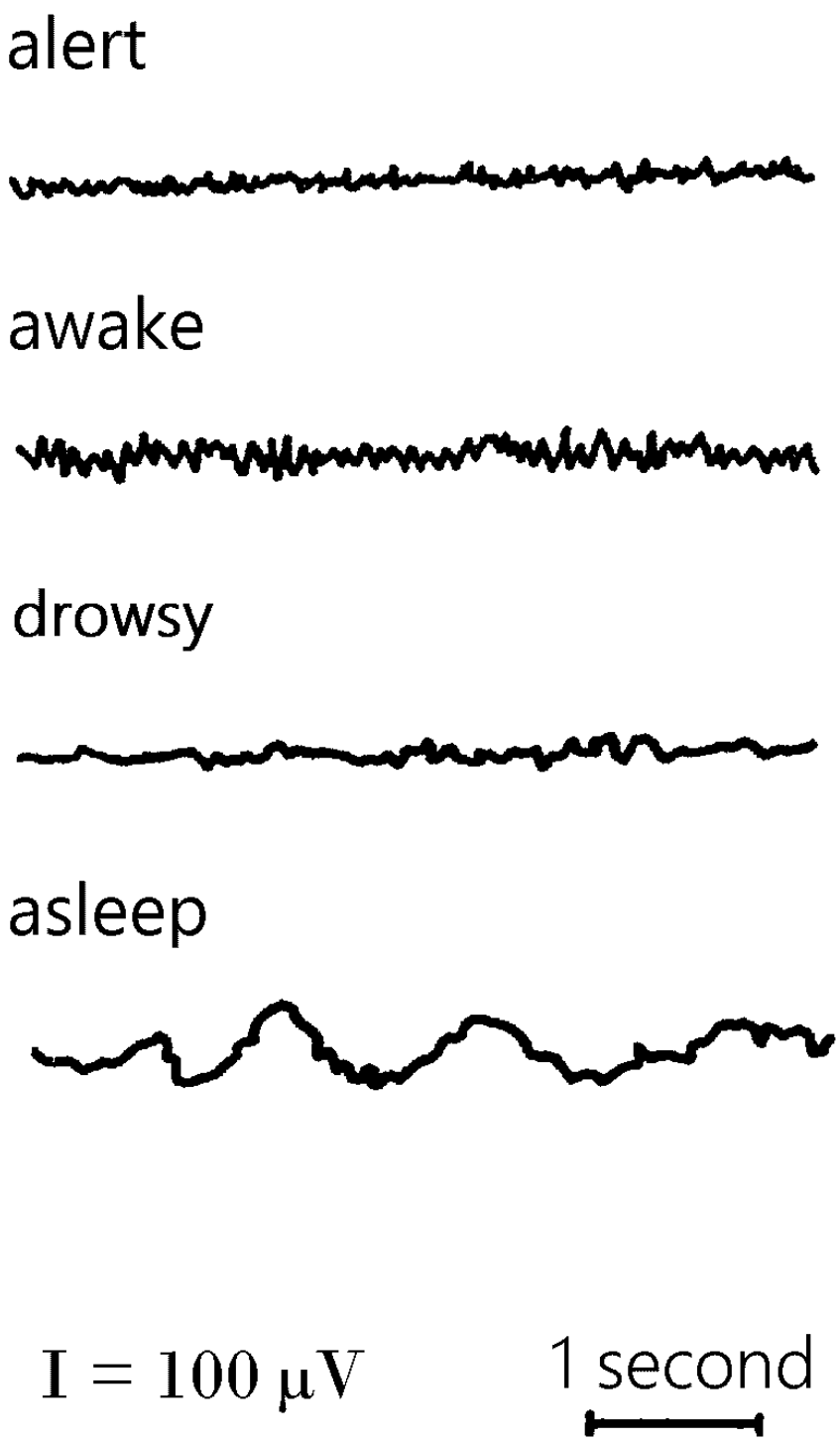


Figure C.35.: Changes in the normal EEG during alertness, drowsiness and sleep, from [18]

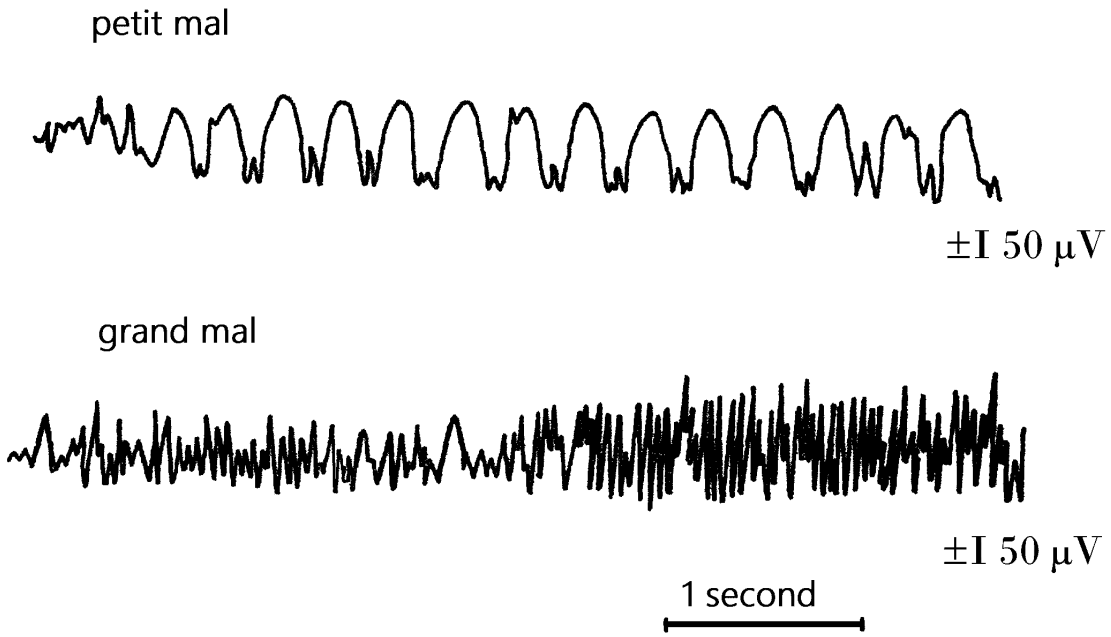


Figure C.36.: Abnormal EEGs

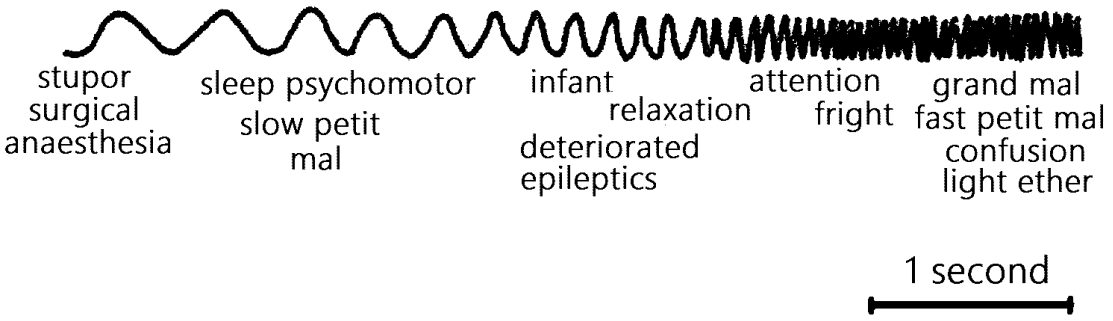


Figure C.37.: The cortical frequency spectrum, from [48]

Figure C.38.: Evoked potential in the visual cortex of a lightly anaesthetized cat, from [18]

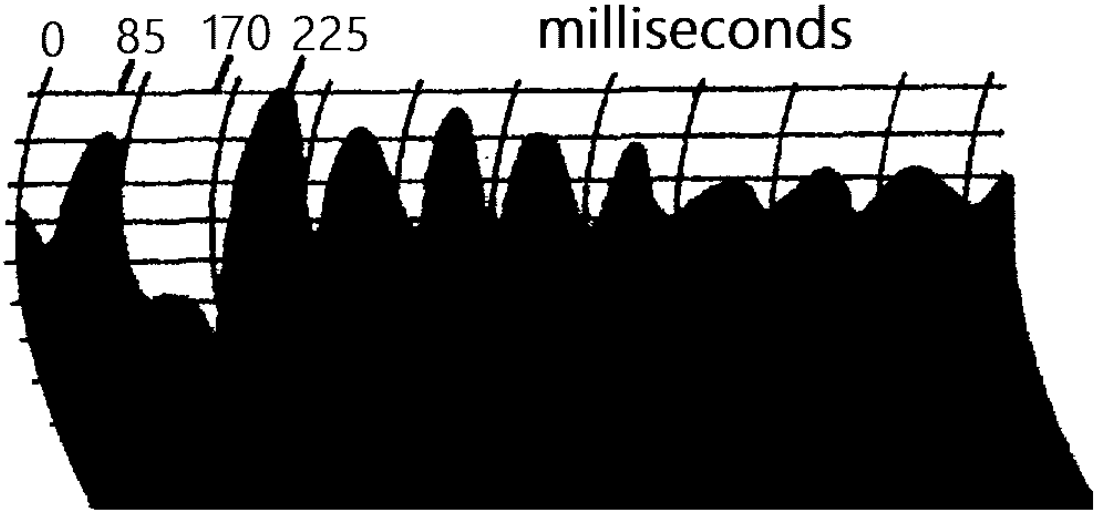


Figure C.39.: Response to flash in man, from [18]

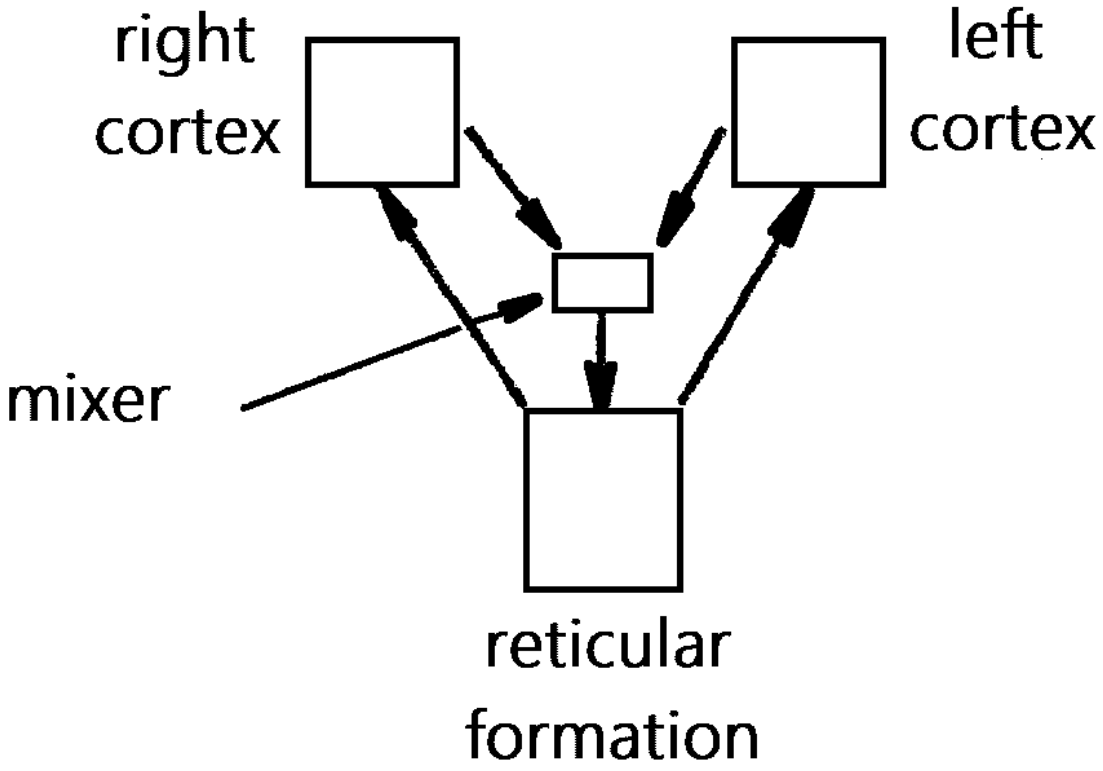


Figure C.40.: Cortico-subcortical circuits, from [88]

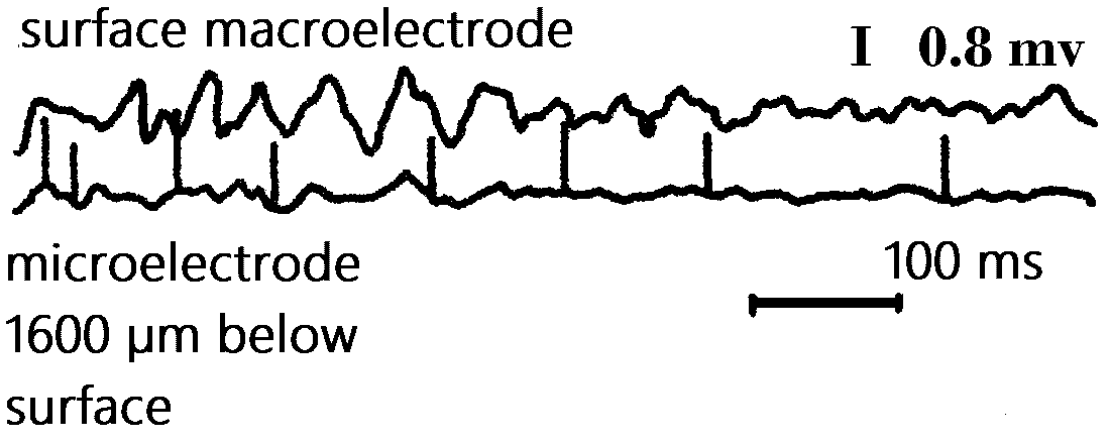


Figure C.41.: Spike patterns and EEG, from [96]

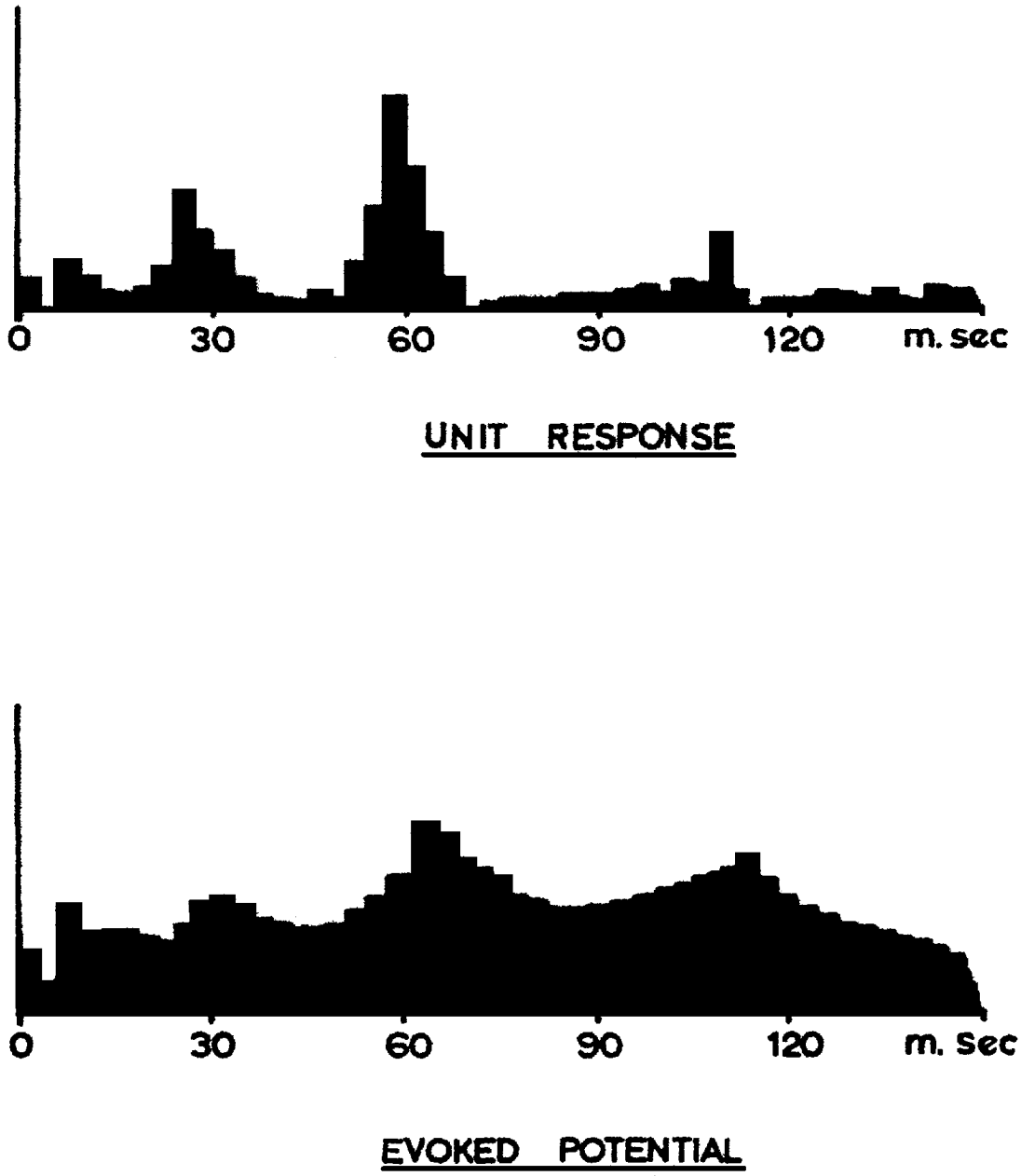


Figure C.42.: Correlation between PSTHs and mean slow potential, [127]

Figure C.43.: "ON", "OFF", and "ON OFF" units

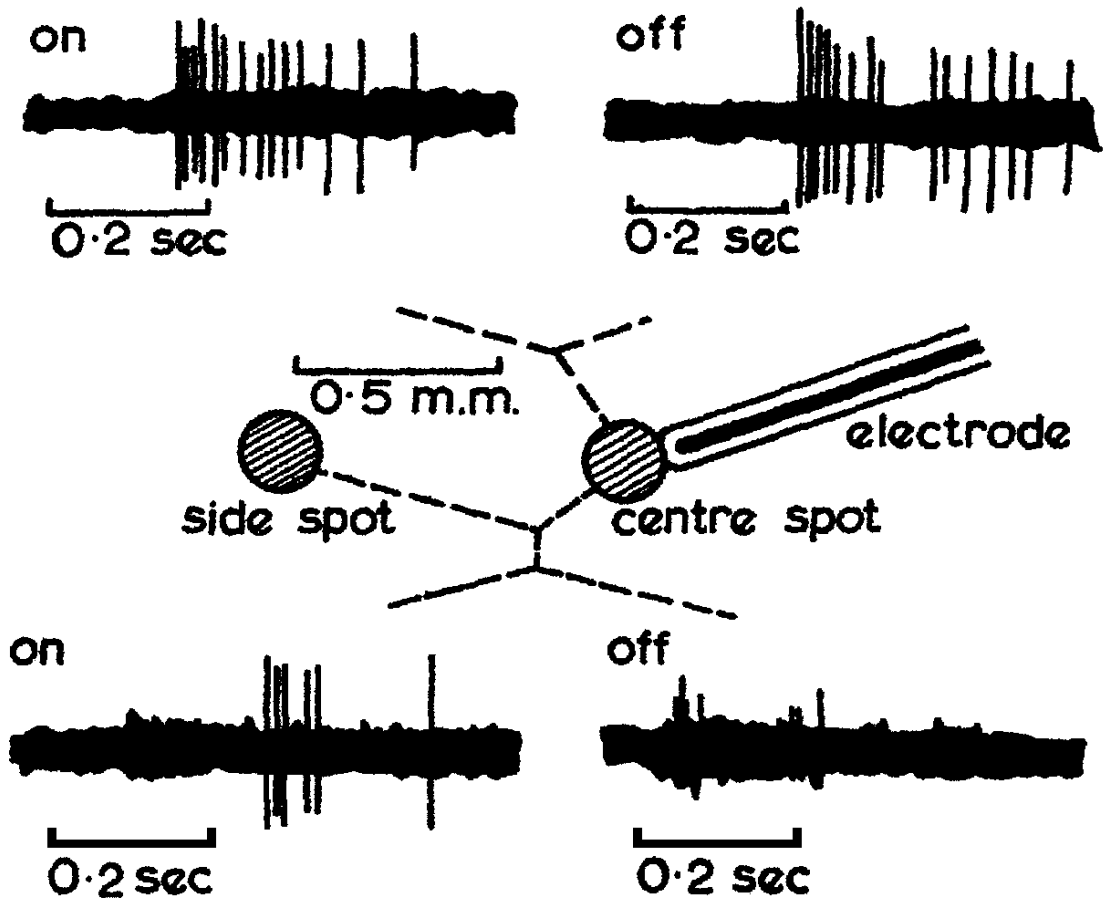


Figure C.44.: Discharge produced by central spot inhibited by side-spot turned on and off with it. top: centre spot alone. bottom: centre and side spots, from [7]

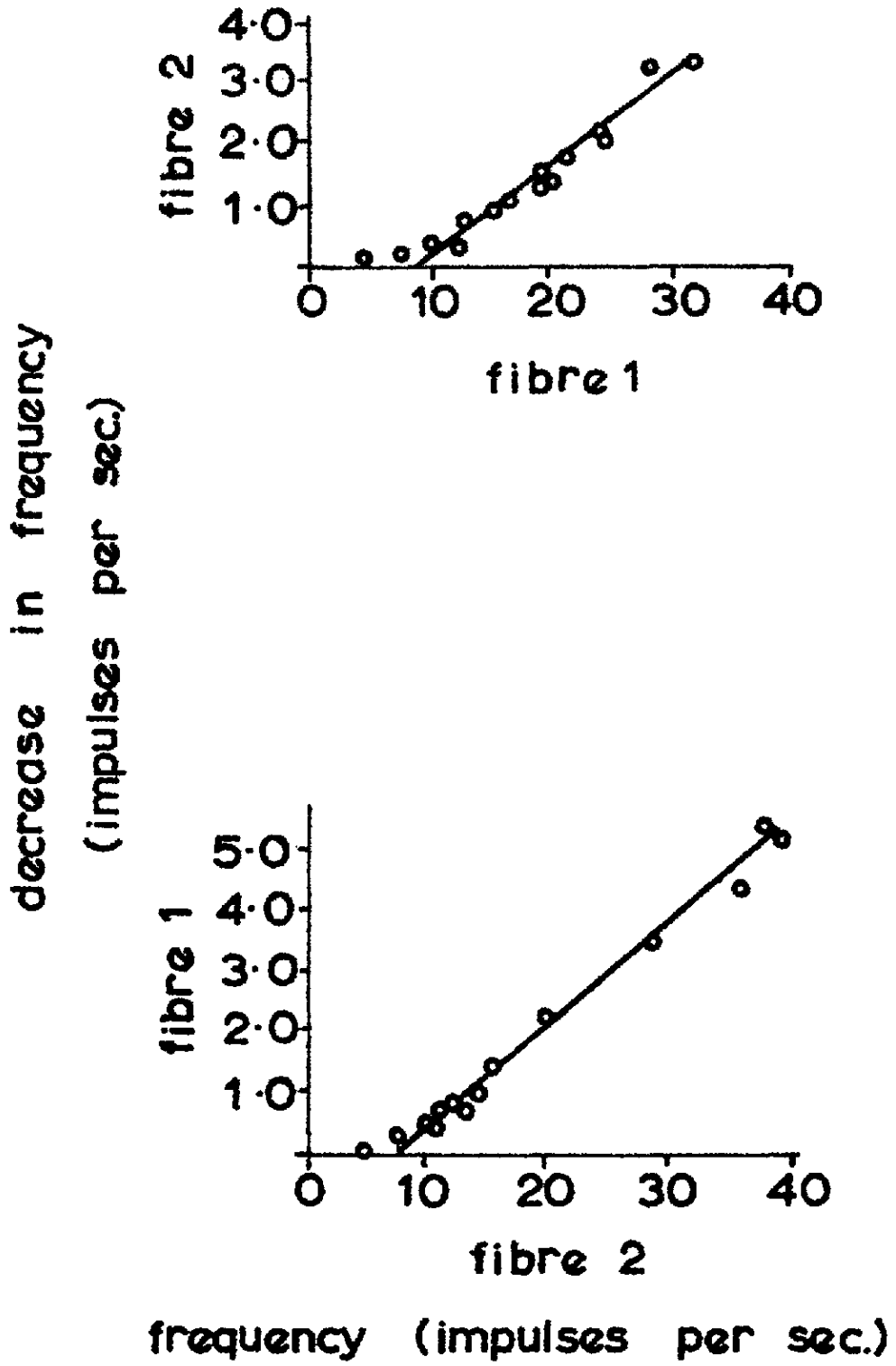


Figure C.45.: Mutual inhibition of two receptor units, from [66]

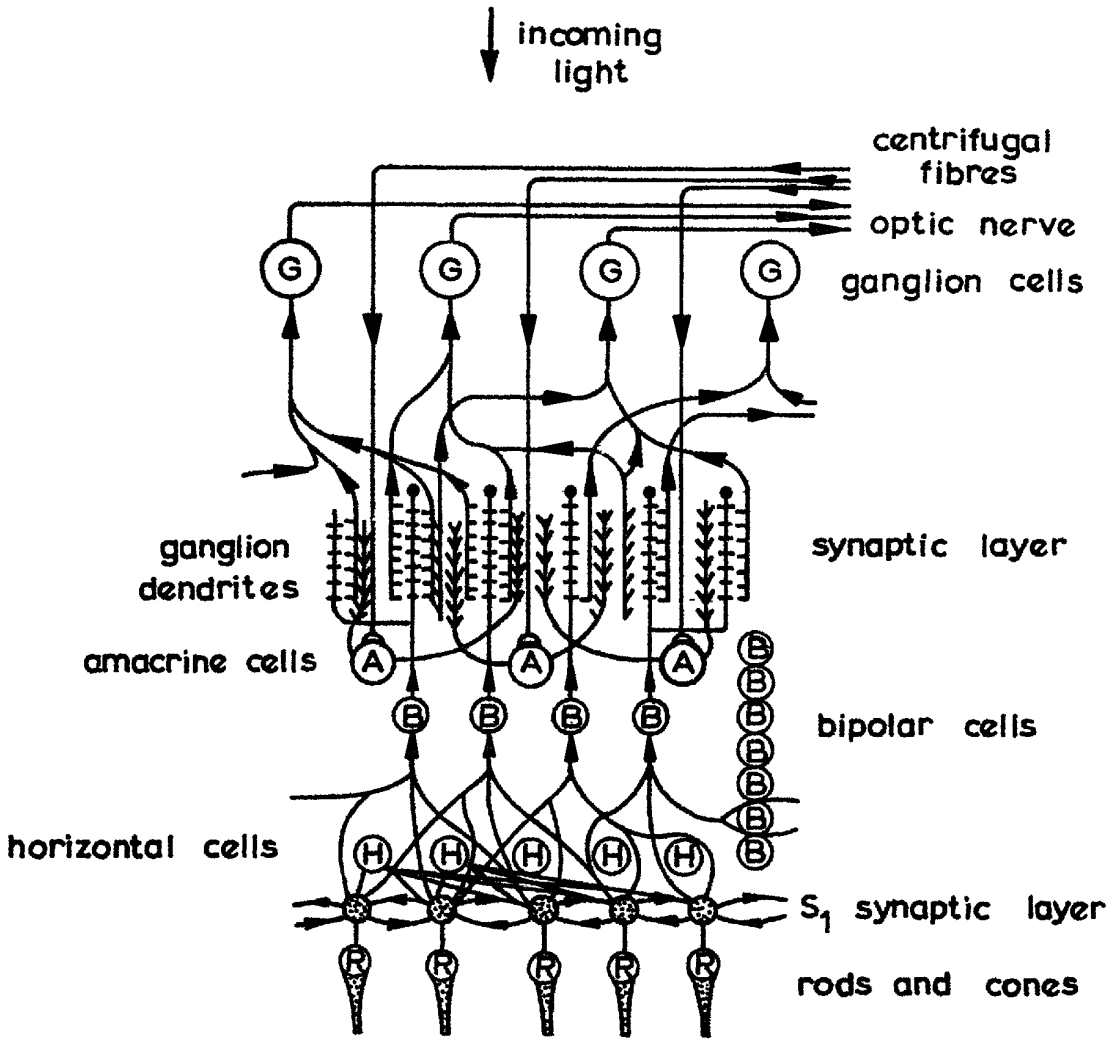


Figure C.46.: Retinal anatomy, from [113]

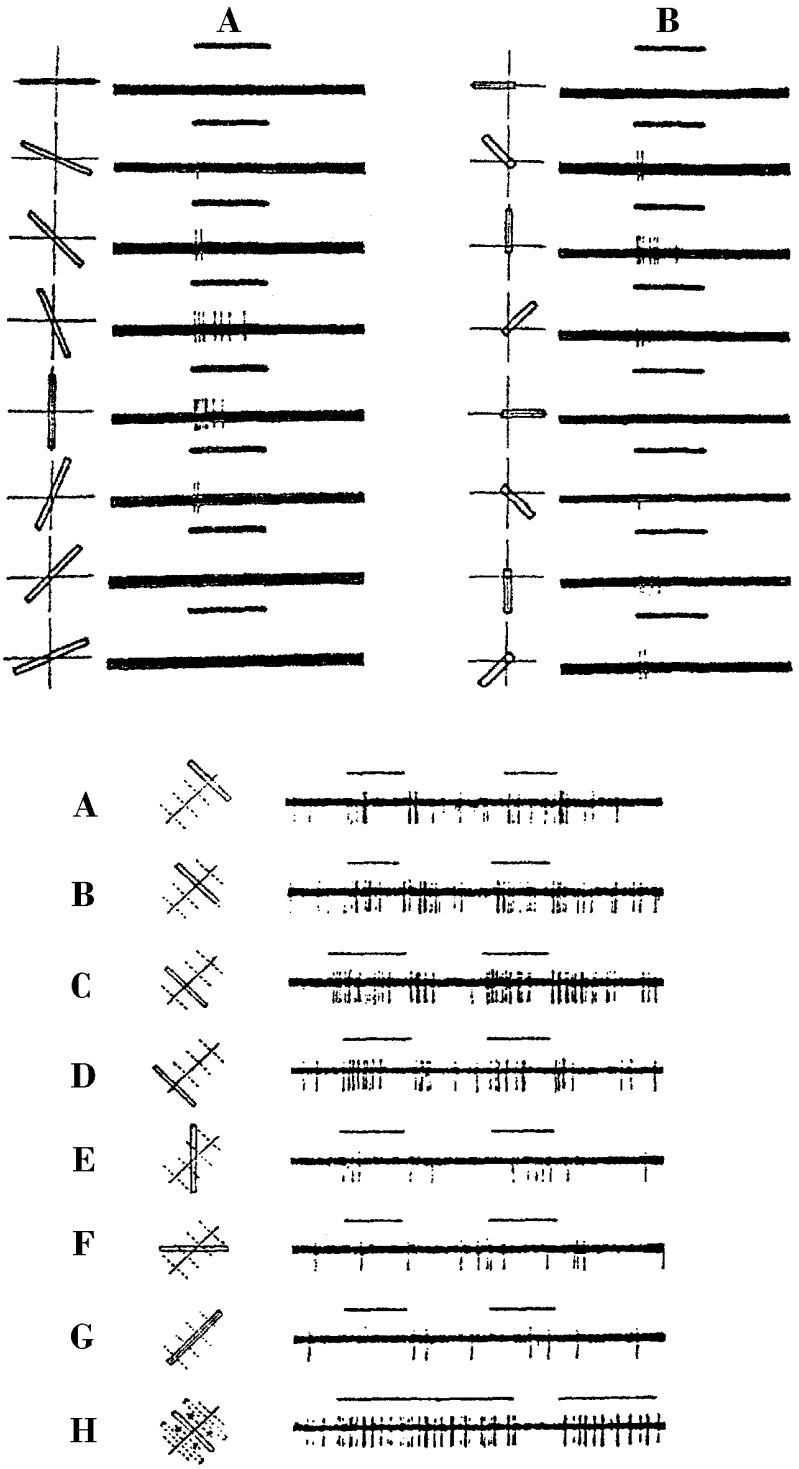


Figure C.47.: A simple and a complex cell, from [72], [73]

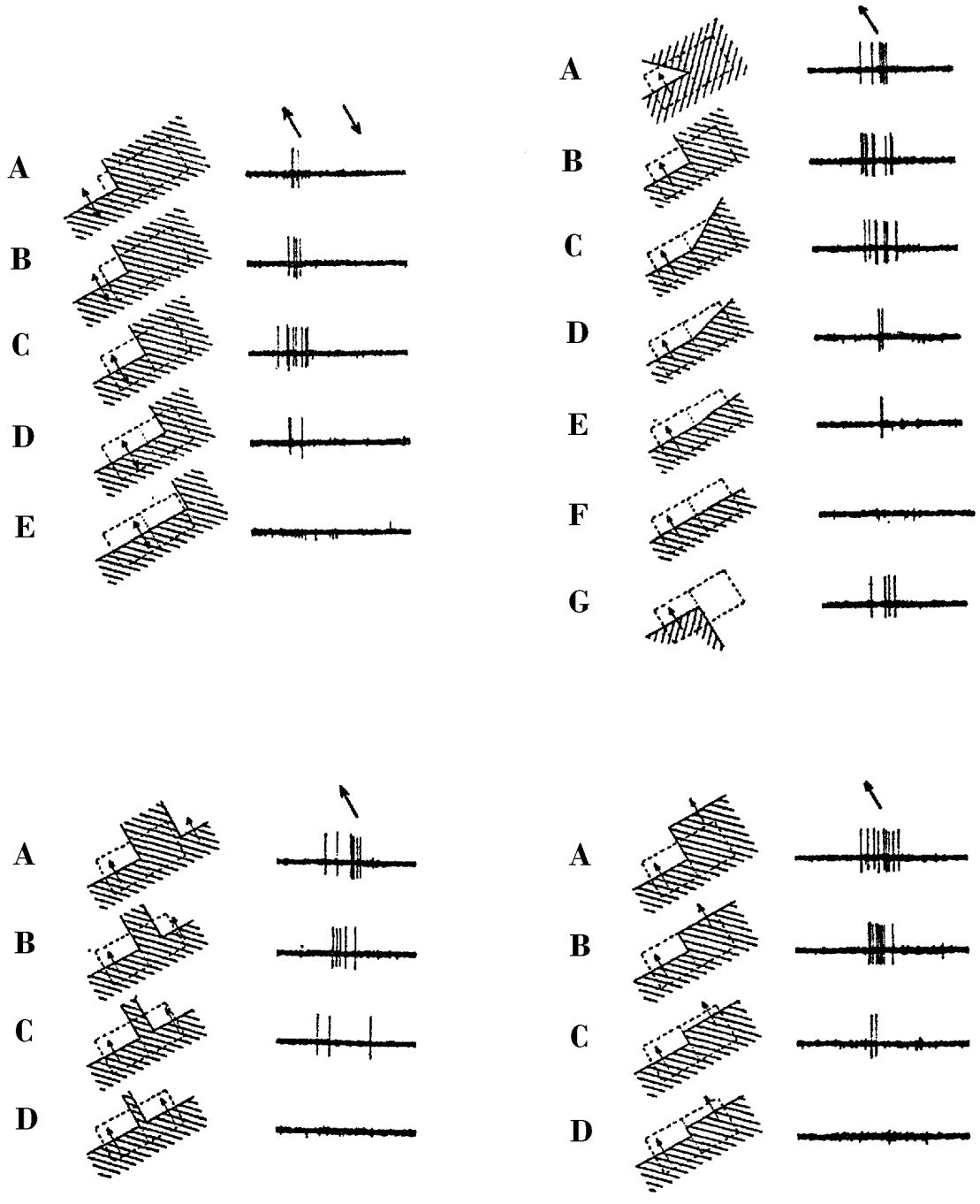


Figure C.48.: A hypercomplex cell, from [74]

Bibliography

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover Publications, 1 edition, 1965.

[2] A. R. Adolph. Spontaneous slow potential fluctuations in the limulus photoreceptor. *Journal of General Physiology*, 48(2):297–322, 1964.

[3] E. D. Adrian. The all-or-none principle in nerve. *The Journal of Physiology*, 47(6):460–474, 1914.

[4] P. Andersen and J. C. Eccles. Inhibitory phasing of neuronal discharge. *Nature*, 196(4855):645–647, 1962.

[5] P. Andersen and T. A. Sears. The role of inhibition in the phasing of spontaneous thalamo-cortical discharge. *The Journal of Physiology*, 173(3):459–480, 1964.

[6] N. T. J. Bailey. *The Mathematical Theory of Epidemics*. London: Charles Griffin & Co., Ltd., 1 edition, 1957.

[7] H. B. Barlow. Summation and inhibition in the frog's retina. *The Journal of Physiology*, 119(1):69–88, 1953.

[8] H. B. Barlow. Increment thresholds at low intensities considered as signal/noise discriminations. *The Journal of Physiology*, 136(3):469–488, 1957.

[9] D. H. Barron and B. H. C. Matthews. The interpretation of potential changes in the spinal cord. *The Journal of Physiology*, 92(3):276–321, 1938.

[10] G. Baumgartner and P. Hakas. Reaktionen einzelner Opticusneurone und corticaler Nervenzellen der Katze im Hell-Dunkel-Grenzfeld (Simultankontrast). *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, 270(1):29–29, 1959.

[11] H. Berger. Über das Elektrenkephalogramm des Menschen. *Archiv für Psychiatrie und Nervenkrankheiten*, 87(1):527–570, 1929.

[12] R. L. Beurle. Properties of a mass of cells capable of regenerating pulses. *Philosophical Transactions of the Royal Society B*, 240(669):55–94, 1956.

[13] G. D. Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences of the United States of America*, 17:656–660, 1931.

[14] E. A. Blair and J. Erlanger. A comparison of the characteristics of axons through their individual electrical responses. *American Journal of Physiology*, 106(3):524–564, 1933.

[15] M. Blum. Properties of a neuron with many inputs. Principles of Self-Organization, pages 95–119. Pergamon Press, New York, 1962.

[16] M. A. B. Brazier. A study of the late response to flash in the cortex of the cat. *Acta physiologica et pharmacologica Neerlandica*, 6(3):692–714, 1957.

[17] M. A. B. Brazier. Long-persisting electrical traces in the brain of man and their possible relationship to higher nervous activity. In H. H. Jasper and G. D. Smirnov, editors, *Electroencephalography of Higher Nervous Activity*, The Moscow Colloquium on Electroencephalography of Higher Nervous Activity and Its Impact on International Brain Research, pages 347–358. Academy of Sciences, U.S.S.R., 1958.

[18] M. A. B. Brazier. *The Electrical Activity of the Nervous System*. Macmillan, 2 edition, 1960.

[19] F. Bremer. Cerebral and cerebellar potentials. *Physiological Reviews*, 38(3):357–388, 1958.

[20] F. Bremer. Physiology of the corpus callosum. In *The Brain and Human Behavior*, volume 36 of *Proceedings of the Association for Research in Nervous and Mental Disease*, pages 424–448. 1958.

[21] B. D. Burns. Some properties of isolated cerebral cortex in the unanaesthetized cat. *The Journal of Physiology*, 112(1-2):156–175, 1951.

[22] B. D. Burns, W. Heron, and R. Pritchard. Physiological excitation of visual cortex in cat's unanaesthetized isolated forebrain. *Journal of Neurophysiology*, 25(2):165–181, 1962.

[23] I. Calma. Thalamo-cortical relations in the sensory nuclei of the cat. *Nature*, 205(4969):394–396, 1965.

[24] H. Caspers. Über die Beziehungen zwischen Dendritenpotential und Gleichspannung an der Hirnrinde. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, 269(2):157–181, 1959.

[25] H. T. Chang. Interaction of evoked cortical potentials. *Journal of Neurophysiology*, 16(2):133–144, 1953.

[26] R. Courant. *Differential and Integral Calculus, 2. Vol.* Blackie & Son Limited, London, 2 edition, 1962.

[27] J. D. Cowan. In preparation.

[28] J. D. Cowan. In A. R. Meetham and R. A. Hudson, editors, *Encyclopaedia of Information, Linguistics and Control*, page 718. Pergamon Press, New York and London, 1969.

[29] D. R. Cox. *Renewal Theory.* London: Methuen & Co. Ltd & New York: John Wiley & Sons Inc., 1 edition, 1962.

[30] B. G. Cragg and H. N. V. Temperley. The organisation of neurones: A co-operative analogy. *Electroencephalography and Clinical Neurophysiology*, 6:85–92, 1954.

[31] K. J. W. Craik. *The Nature of Explanation.* Cambridge University Press, 1 edition, 1943.

[32] O. D. Creutzfeldt. Some problems of information transmission in the visual system. In J. C. Eccles, editor, *Brain and Conscious Experience*, pages 138–164. Berlin-Heidelberg-New York: Springer, 1 edition, 1964.

[33] E. W. Dempsey and R. S. Morison. Some afferent diencephalic pathways related to cortical potentials in the cat. *American Journal of Physiology*, 131(3):718–731, 1940.

[34] E. W. Dempsey and R. S. Morison. The production of rhythmically recurrent cortical potentials after localized thalamic stimulation. *American Journal of Physiology*, 135(2):293–300, 1941.

[35] J. C. Eccles. *Physiology of Nerve Cells.* The Johns Hopkins Press, 1 edition, 1957.

[36] J. C. Eccles. *The Physiology of Synapses.* Berlin-Heidelberg-New York: Springer, 1 edition, 1964.

[37] J. C. Eccles, P. Fatt, and K. Koketsu. Cholinergic and inhibitory synapses in a pathway from motor-axon collaterals to motoneurones. *The Journal of Physiology*, 126(3):524–562, 1954.

[38] P. Elias. Computation in the presence of noise. *IBM Journal of Research and Development*, 2:346–353, 1958.

[39] B. G. Farley. A neural network model and the "slow potentials" of electrophysiology. In R. W. Stacy and B. D. Waxman, editors, *Computers in Biomedical Research. 2 Vols.*, pages 265–294. Academic Press Inc., New York, 1 edition, 1965.

[40] P. Fatt and B. Katz. Spontaneous subthreshold activity at motor nerve endings. *The Journal of Physiology*, 117(1):109–128, 1952.

[41] W. Feller. *An Introduction to Probability Theory and Its Applications*. John Wiley & Sons, Inc., New York, 2 edition, 1968.

[42] A. E. Fessard. Information processing in the nervous system. volume 3 of *Proceedings of the International Union of Physiological Sciences*, pages 415–418. Amsterdam, New York, Excerpta Medica Foundation, 1964.

[43] J. Field, editor. *Handbook of Physiology, Section 1: Neurophysiology*. American Physiological Society, Washington D.C., 1 edition, 1959.

[44] R. Fitzhugh. Thresholds and plateaus in the Hodgkin-Huxley nerve equations. *Journal of General Physiology*, 43(5):867–896, 1960.

[45] S. S. Fox and J. H. O'Brien. Duplication of evoked potential waveform by curve of probability of firing of a single cell. *Science*, 147(3660):888–890, 1965.

[46] C. D. Geisler and J. M. Goldberg. A stochastic model of the repetitive activity of neurons. *Biophysical Journal*, 6(1):53–69, 1966.

[47] G. L. Gerstein and B. Mandelbrot. Random walk models for the spike activity of a single neuron. *Biophysical Journal*, 4(1):41–68, 1964.

[48] F. A. Gibbs and E. L. Gibbs. *Atlas of Electroencephalography*. Addison-Wesley Press, 1958.

[49] J. W. Gibbs. *Elementary Principles in Statistical Mechanics*. Yale University Press, 1 edition, 1902.

[50] H. Goldstein. *Classical Mechanics*. Oliver and Boyd, 2 edition, 1957.

[51] B. C. Goodwin. *Temporal Organization in Cells. A Dynamic Theory of Cellular Control Processes*. Academic Press Inc., New York, 1 edition, 1963.

[52] B. Grafstein. Mechanism of spreading cortical depression. *Journal of Neurophysiology*, 19(2):154–171, 1956.

[53] C. H. Graham and H. K. Hartline. The response of single visual sense cells to lights of different wave lengths. *Journal of General Physiology*, 18:917–931, 1935.

[54] R. Granit. *Sensory Mechanisms of the Retina*. Oxford University Press, 1 edition, 1947.

[55] R. Granit, D. Kernell, and G. K. Shortess. The behaviour of mammalian motoneurones during long-lasting orthodromic, antidromic and trans-membrane stimulation. *The Journal of Physiology*, 169(4):743–754, 1963.

[56] R. Granit and B. Renkin. Net depolarization and discharge rate of motoneurones, as measured by recurrent inhibition. *The Journal of Physiology*, 158(3):461–475, 1961.

[57] J. D. Green and A. Arduini. Hippocampal electrical activity in arousal. *Journal of Neurophysiology*, 17(6):533–557, 1954.

[58] R. L. Gregory. Models and the localization of function in the central nervous system. In *Mechanization of Thought Processes*, volume 2 of *National Physical Laboratory Symposium No. 10*, pages 566–583. London: Her Majesty's Stationery Office, 1958.

[59] J. S. Griffith. A field theory of neural nets: I: Derivation of field equations. *The Bulletin of Mathematical Biophysics*, 25(1):111–120, 1963.

[60] J. S. Griffith. On the stability of brain-like structures. *Biophysical Journal*, 3(4):299–308, 1963.

[61] J. S. Griffith. A field theory of neural nets: II. properties of the field equations. *The Bulletin of Mathematical Biophysics*, 27(1):187–195, 1965.

[62] D. Haar ter. Theory and applications of the density matrix. *Reports on Progress in Physics*, 24(1):304–362, 1961.

[63] S. Hagiwara. Analysis of interval fluctuation of the sensory nerve impulse. *Japanese Journal of Physiology*, 4:234–240, 1954.

[64] H. K. Hartline. The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *American Journal of Physiology*, 121(2):400–415, 1938.

[65] H. K. Hartline. The receptive fields of optic nerve fibers. *American Journal of Physiology*, 130(4):690–699, 1940.

[66] H. K. Hartline and F. Ratliff. Inhibitory interaction of receptor units in the eye of limulus. *Journal of General Physiology*, 40(3):357–376, 1957.

[67] D. O. Hebb. *The Organization of Behavior: A Neuropsychological Theory*. John Wiley & Sons, Inc., New York, 1 edition, 1949.

[68] A. L. Hodgkin. The local electric changes associated with repetitive action in a non-medullated axon. *The Journal of Physiology*, 107(2):163–181, 1948.

[69] A. L. Hodgkin and A. D. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4):500–544, 1952.

[70] M. Hoopen ten. Mathematical model of cooperative processes in neural populations. In E. R. Caianiello, editor, *Cybernetics of Neural Processes*, volume 31 of *Quaderni de La Ricerca scientifica*, pages 279–290. Consiglio nazionale delle ricerche, 1965.

[71] M. Hoopen Ten. Probabilistic firing of neurons considered as a first passage problem. *Biophysical Journal*, 6(4):435–451, 1966.

[72] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148(3):574–591, 1959.

[73] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1):106–154, 1962.

[74] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 28(2):229–289, 1965.

[75] A. Hugelin, S. Dumont, and N. Paillas. Formation réticulaire et transmission des informations auditives au niveau de l'oreille moyenne et des voies acoustiques centrales. *Electroencephalography and Clinical Neurophysiology*, 12(4):797–818, 1960.

[76] E. R. John, D. S. Ruchkin, and J. Villegas. Experimental background: Signal analysis and behavioral correlates of evoked potential configurations in cats. *Annals of the New York Academy of Sciences*, 112(1):362–420, 1964.

[77] E. R. Kandel, W. A. Spencer, and F. J. Brinley. Electrophysiology of hippocampal neurons: I. sequential invasion and synaptic organization. *Journal of Neurophysiology*, 24(3):225–242, 1961.

[78] Y. Katsuki, T. Sumi, H. Uchiyama, and T. Watanabe. Electric responses of auditory neurons in cat to sound stimulation. *Journal of Neurophysiology*, 21(6):569–588, 1958.

[79] M. Kendall and A. Stuart. *The advanced theory of statistics, vol. 1.* London: Charles Griffin & Co., Ltd., 2 edition, 1963.

[80] D. Kernell. High-frequency repetitive firing of cat lumbosacral motoneurones stimulated by long-lasting injected currents. *Acta Physiologica Scandinavica*, 65(1-2):74–86, 1965.

[81] D. Kernell. The limits of firing frequency in cat lumbosacral motoneurones possessing different time course of afterhyperpolarization. *Acta Physiologica Scandinavica*, 65(1-2):87–100, 1965.

[82] E. H. Kerner. A statistical mechanics of interacting biological species. *The Bulletin of Mathematical Biophysics*, 19(2):121–146, 1957.

[83] E. H. Kerner. Further considerations on the statistical mechanics of biological associations. *The Bulletin of Mathematical Biophysics*, 21(2):217–255, 1959.

[84] E. H. Kerner. On the volterra-lotka principle. *The Bulletin of Mathematical Biophysics*, 23(2):141–157, 1961.

[85] N. Y.-S. Kiang. *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve*. M.I.T. Press, 1 edition, 1966.

[86] M. R. Klee, K. Offenloch, and J. Tigges. Cross-correlation analysis of electroencephalographic potentials and slow membrane transients. *Science*, 147(3657):514–521, 1965.

[87] S.C. Kleene. Representation of events in nerve nets and finite automata. In C. E. Shannon and J. McCarthy, editors, *Automata Studies*, Annals of Mathematics Studies, pages 3–42. Princeton University Press, 1956.

[88] W. P. Koella and A. Ferry. Cortico-subcortical homeostasis in the cat's brain. *Science*, 142(3592):586–589, 1963.

[89] N. Krylov and N. Bogolyubov. La théorie générale de la mesure dans son application à l'étude des systèmes dynamiques de la mécanique non linéaire. *Annals of Mathematics*, 38(1):65–113, 1937.

[90] S. W. Kuffler. Neurons in the retina: organization, inhibition and excitation problems. In *The Neuron*, volume 17 of *Cold Spring Harbor Symposia on Quantitative Biology*, pages 281–292, 1952.

[91] S. W. Kuffler, R. Fitzhugh, and H. B. Barlow. Maintained activity in the cat's retina in light and darkness. *Journal of General Physiology*, 40(5):683–702, 1957.

[92] M. Kuno. Quantal components of excitatory synaptic potentials in spinal motoneurones. *The Journal of Physiology*, 175(1):81–99, 1964.

[93] K. S. Lashley. In search of the engram. In *Symposium of the society for experimental biology*, volume 4, pages 454–455, 468–473, 477–480. Cambridge University Press, 1950.

[94] A. A. P. Leao. Spreading depression of activity in the cerebral cortex. *Journal of Neurophysiology*, 7(6):359–390, 1944.

[95] J. Y. Lettvin, H. R. Maturana, W. S. McCulloch, and W. H. Pitts. What the frog's eye tells the frog's brain. *Proceedings of the IRE (Institute of Radio Engineers)*, 47(11):1940–1951, 1959.

[96] C.-L. Li and H. Jasper. Microelectrode studies of the electrical activity of the cerebral cortex in the cat. *The Journal of Physiology*, 121(1):117–140, 1953.

[97] J. C. Lilly. Forms and figures in the electrical activity seen in the surface of the cerebral cortex. In *The Biology of Mental Health and Disease*, pages 206–219. New York: Hoeber-Harper, 1 edition, 1952. The twenty-seventh annual conference of the Milbank Memorial Fund.

[98] W. R. Loewenstein. Excitation and inactivation in a receptor membrane. *Annals of the New York Academy of Sciences*, 94(2):510–534, 1961.

[99] R. Lorente de Nó. Cerebral cortex: Architecture, intracortical connections, motor projections. In J. F. Fulton, editor, *Physiology of the Nervous System*, Outlines of Physiology Series. Oxford University Press, 3 edition, 1949.

[100] A. J. Lotka. *Elements of Physical Biology*. Williams and Wilkins, Baltimore, 1 edition, 1925.

[101] C. A. Marsan and A. Morillo. Callosal and specific response in the visual cortex of cat. *Archives Italiennes De Biologie*, 101(1):1–29, 1963.

[102] W. McCulloch. The stability of biological systems. In *Homeostatic Mechanisms*, volume 10 of *Brookhaven Symposia in Biology*, pages 207–15. Brookhaven National Lab., Upton, N.Y., 1957.

[103] W. S. McCulloch, C. Graf, and H. W. Magoun. A cortico-bulbo-reticular pathway from area 4-s. *Journal of Neurophysiology*, 9(2):127–, 1946.

[104] W. S. McCulloch and W. H. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5(4):115–133, 1943.

[105] E. F. Moore and C. E. Shannon. Reliable circuits using less reliable relays. *Journal of the Franklin Institute*, 262(3):191–208, 1956.

[106] P. A. P. Moran. *Statistical Processes of Evolutionary Theory*. Oxford University Press, 1 edition, 1962.

[107] G. Moruzzi and H. W. Magoun. Brain stem reticular formation and activation of the EEG. *Electroencephalography and Clinical Neurophysiology*, 1(1-4):455–473, 1949.

[108] V. B. Mountcastle. Modality and topographic properties of single neurons of cat's somatic sensory cortex. *Journal of Neurophysiology*, 20(4):408–434, 1957.

[109] J. Neumann von. Probabilistic logics and the synthesis of reliable organisms from unreliable components. In C. E. Shannon and J. McCarthy, editors, *Automata Studies*, Annals of Mathematics Studies, pages 43–98. Princeton University Press, 1956.

[110] J. L. O'Leary and S. Goldring. Changes associated with forebrain excitation processes: D.C. potentials of the cerebral cortex. In J. Field, editor, *Handbook of Physiology, Section 1: Neurophysiology*, pages 315–328. American Physiological Society, Washington D.C., 1 edition, 1959.

[111] R. Pearl. *The Biology of Death*. Lippincott, Philadelphia, 1 edition, 1922.

[112] C. Pecher. La fluctuation d'excitabilité de la fibre nerveuse. *Archives Internationales de Physiologie et de Biochimie*, 49(2):129–152, 1939.

[113] C. Pedler. Rods and cones – a fresh approach. In A. V. S. DeReuck and J. Knight, editors, *Colourvision – Physiology and Experimental Psychology*, Ciba Foundation Symposium, pages 52–88. J. & A. Churchill, Ltd., London, W.I., 1965.

[114] W. W. Peterson. *Error-Correcting Codes*. M.I.T. Press, 1 edition, 1961.

[115] C. G. Phillips. Actions of antidromic pyramidal volleys on single betz cells in the cat. *Quarterly Journal of Experimental Physiology and Cognate Medical Sciences*, 44(1):1–25, 1959.

[116] W. H. Pierce. Adaptive vote-takers improve the use of redundancy. In R. H. Wilcox and W. C. Mann, editors, *Redundancy Techniques in Computing Systems*, pages 229–250. Spartan Books, Washington, D. C., 1 edition, 1962.

[117] G. F. Poggio and L. J. Viernstein. Time series analysis of impulse sequences of thalamic somatic sensory neurons. *Journal of Neurophysiology*, 27(4):517–545, 1964.

[118] S. L. Polyak. *The verterebrate visual system*. The University of Chicago Press, 1 edition, 1958.

[119] I. Prigogine. *Introduction to Thermodynamics of Irreversible Processes*. John Wiley & Sons, Inc., New York, 3 edition, 1967.

[120] W. Rall. Branching dendritic trees and motoneuron membrane resistivity. *Experimental Neurology*, 1(5):491–527, 1959.

[121] W. Rall. Theoretical significance of dendritic trees for neuronal input-output relations. In R. F. Reiss, editor, *Neural Theory and Modelling*, pages 73–97. Stanford University Press, 1 edition, 1964.

[122] W. Rall and C. C. Hunt. Analysis of reflex variability in terms of partially correlated excitability fluctuation in a population of motoneurons. *Journal of General Physiology*, 39(3):397–422, 1956.

[123] S. Ramón-Y.-Cajal. *Les Nouvelles Idées sur la Structure du Système Nerveux Chez l'Homme Et Chez les Vertébrés*. Paris, C. Reinwald & Cie, 1894.

[124] S. L. Ranson and S. W. Clark. *The Anatomy of the Nervous System – Its Development and Function*. W. B. Saunders Company, 10 edition, 1959.

[125] B. Renshaw. Influence of discharge of motoneurons upon excitation of neighboring motoneurons. *Journal of Neurophysiology*, 4(2):167–183, 1941.

[126] B. Renshaw, A. Forbes, and B. R. Monson. Activity of isocortex and hippocampus: electrical studies with micro-electrodes. *Journal of Neurophysiology*, 3(1):74–105, 1940.

[127] A. D. J. Robertson. In preparation.

[128] A. D. J. Robertson. Correlation between unit activity and slow potential changes in the unanaesthetized cerebral cortex of the cat. *Nature*, 208(5012):757–758, 1965.

[129] A. D. J. Robertson, 1967. Personal Communication.

[130] R. W. Rodieck, N. Y.-S. Kiang, and G. L. Gerstein. Some quantitative methods for the study of spontaneous activity of single neurons. *Biophysical Journal*, 2(4):351–, 1962.

[131] F. Rosenblatt. The perceptron: A theory of statistical separability in cognitive systems. Technical Report VG-1196-G-1, Cornell Aeronautical Laboratory, 1958.

[132] F. Rosenblatt. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Spartan Books, Washington DC, 1 edition, 1962.

[133] W. A. Rosenblith. Electrical responses from the auditory nervous system. In *Proceedings of the Symposium on Information Networks*, volume 3 of *Microwave Research Institute symposia series*, pages 223–247, 1954.

[134] W. A. Rosenblith, editor. *Processing Neuroelectric Data*. M.I.T. Press, 1 edition, 1959.

[135] W. A. Rosenblith. *Sensory Communication*. M.I.T. Press, 1 edition, 1961.

[136] G. F. Rossi and A. Zanchetti. The brain stem reticular formation: Anatomy and physiology. *Archives Italiennes de Biologie*, 95(3-4):199–438, 1957.

[137] M. G. Saunders. Amplitude probability density sttidies of alpha activity in the EEG. In Medical Research Council, editor, *Mathematics and Computer Science in Biology and Medicine*, Proceedings of the Conference held by the Medical Research Council in association with the Health Departments: Oxford, page 317. London: Her Majesty's Stationery Office, 1965.

[138] A. J. Sefton and W. Burke. Reverberator inhibitory circuits in the lateral geniculate nucleus of the rat. *Nature*, 205(4978):1325–1326, 1965.

[139] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[140] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(4):623–656, 1948.

[141] C. S. Sherrington. *Man on His Nature*. Cambridge University Press, 1 edition, 1940.

[142] D. A. Sholl. *The organization of the cerebral cortex*. John Wiley & Sons, Inc., New York, 1 edition, 1956.

[143] L. J. Slater. *Confluent Hypergeometric Functions*. Cambridge University Press, 1 edition, 1960.

[144] D. R. Smith and C. H. Davidson. Maintained activity in neural nets. *Journal of the Association for Computing Machinery*, 9(2):268–279, 1962.

[145] G. K. Smith and D. R. Smith. A statistical analysis of the continual activity of single cortical neurones in the cat unanaesthetized isolated forebrain. *Biophysical Journal*, 5(1):47–74, 1965.

[146] R. W. Sperry. Orderly function with disordered structure. Principles of Self-Organization, pages 279–291. Pergamon Press, New York, 1961.

[147] R. B. Stein. A theoretical analysis of neuronal variability. *Biophysical Journal*, 5(2):173–194, 1965.

[148] K. Suhara and S. Uemura. A magnetic type short range EEG analyzer and its application, statistical analysis of EEG by the frequency analyzer and the analog correlator. In F. H. Bostem, editor, *Medical Electronics*, Proceedings of the fifth International Conference on Medical Electronics, pages 736–752. Institute of Electrical Engineers, 1963.

[149] W. K. Taylor. Electric simulation of some nervous systems functional activities. In E. C. Cherry, editor, *Information Theory*, volume 3 of *3rd London Symposium on Information Theory*, pages 314–328. Butterworths, London, 1956.

[150] A. C. Terzuolo and Bo E. Gernandt. Spinal unit activity during synchronization of a convulsive type (strychnine tetanus). *American Journal of Physiology*, 186(2):263–270, 1956.

[151] A. M. Turing. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, s2–42(1):230–265, 1937.

[152] A. M. Uttley. The design of conditional probability computers. *Information and Control*, 2(1):1–24, 1959.

[153] A. M. Uttley. The transmission of information and the effect of local feedback in theoretical and neural networks. *Brain Research*, 2(1):21–50, 1966.

[154] A. A. Verveen. On the fluctuation of threshold of the nerve fibre. In D. B. Tower and J. P. Schade, editors, *Structure And Function of the Cerebral Cortex*, pages 282–286. Elsevier Science Ltd, 1 edition, 1960.

[155] A. A. Verveen. *Fluctuation in excitability. Research report on signal transmission in nerve fibers.* PhD thesis, University of Amsterdam, Central Institute for Brain Research, 1961.

[156] A. A. Verveen. Axon diameter and fluctuation in excitability. *Acta morphologica Neerlando-Scandinavica*, 5:79–85, 1962.

[157] A. A. Verveen and H. E. Derksen. Fluctuations in membrane potential of axons and the problem of coding. *Kybernetik*, 2:152–160, 1965.

[158] A. A. Verveen and H. E. Derksen. Fluctuations of resting neural membrane potential. *Science*, 151(3716):1388–1389, 1966.

[159] M. Verzeano and K. Negishi. Neuronal activity in cortical and thalamic networks: A study with multiple microelectrodes. *Journal of General Physiology*, 43(6):177–195, 1960.

[160] L. J. Viernstein and R. G. Grossman. Neural discharge patterns in the transmission of sensory information. In E. C. Cherry, editor, *Information Theory*, volume 4 of *4th London Symposium on Information Theory*, pages 252–269. Butterworths, London, 1961.

[161] D. O. Walter, J. M. Rhodes, D. Brown, and W. R. Adey. Comprehensive spectral analysis of human EEG generators in posterior cerebral regions. *Electroencephalography and Clinical Neurophysiology*, 20(3):224–237, 1966.

[162] W. G. Walter. Slow potential waves in the human brain associated with expectancy, attention and decision. *Archiv für Psychiatrie und Nervenkrankheiten*, 206(3):309–322, 1964.

[163] T. F. Weiss. Quantitative methods in the analysis of neuroelectric activity. In C. N. Smyth, editor, *Medical Electronics*, Proceedings of the Second International Conference on Medical Electronics, pages 148–154. Institute of Electrical Engineers, 1960.

[164] E. T. Whitakker. *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*. Cambridge University Press, 1 edition, 1904.

[165] N. Wiener. *Cybernetics; or control and communication in the animal and the machine*. John Wiley & Sons, Inc., New York, 1 edition, 1948.

[166] N. Wiener. Time, communication, and the nervous system. *Annals of the New York Academy of Sciences*, 50(4):197–220, 1948.

[167] N. Wiener. *Nonlinear Problems in Random Theory*. M.I.T. Press, 1 edition, 1958.

[168] N. Wiener and P. J. Schade. *Cybernetics of the Nervous system*, volume 17 of *Progress in Brain Research*. Elsevier Science, 1965.

[169] S Winograd. Coding for logical operations. *IBM Journal of Research and Development*, 6:430–436, 1962.

[170] S. Winograd and J. D. Cowan. *Reliable Computation in the Presence of Noise*. M.I.T. Press, 1 edition, 1963.

[171] A. C. Young. Neural control of respiration. In T. C. Ruch and J. F. Fulton, editors, *Medical Physiology and Biophysics*, pages 813–828. Philadelphia: W. B. Saunders Co., 1960. 18th edition of Howell's Textbook of Physiology.

[172] J. Z. young. *A Model of the Brain*. Oxford University Press, 1 edition, 1964.